

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 11:11:46 ; Search time 55 Seconds

(without alignments)
693.526 Million cell updates/sec

Title: US-10-025-170-2

Perfect score: 686

Sequence: 1 MKAFSPVRSVRKNSLSHSL.....LQASEFFSELMNSDKALCG 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	670.5	97.7	134	5	Aae25009 Human Id2
2	661.5	96.4	134	3	Aab19323 Amino aci
3	661.5	96.4	176	3	Aab56969 Human pro
4	654.5	95.4	134	3	Aab19324 Amino aci
5	654.5	95.4	134	7	Add47773 Rat Prote
6	654.5	95.4	134	7	Add47771 Rat Prote
7	646.5	94.2	134	3	Aab19317 Amino aci
8	640	93.3	133	3	Aab19326 Amino aci
9	522	76.1	135	3	Aab19325 Amino aci
10	297	43.3	59	2	Aay17266 HLH domai
11	266.5	38.8	118	3	Aab19320 Amino aci
12	252.5	36.8	128	3	Aab19321 Amino aci
13	249	36.3	229	5	Abp41248 Human ova
14	242.5	35.3	154	2	Aar53710 Human Id-
15	242.5	35.3	154	5	Aau10351 Human inh
16	242	35.3	155	4	Aau28063 Novel hum
17	241	35.1	161	3	Aab19319 Amino aci
18	240	35.0	251	4	Aau31023 Novel hum
19	236.5	34.5	111	3	Aab59014 Breast an
20	235.5	34.3	148	3	Aab19316 Amino aci
21	226.5	33.0	149	5	Aau10352 Human inh
22	225.5	32.9	164	7	Ades7173 Rat Prote
23	225.5	32.9	164	7	Ades7170 Rat Prote
24	225.5	32.9	164	7	Add45453 Rat Prote
25	224	32.7	148	2	Aar53711 Human Id-

ALIGNMENTS

RESULT 1

AAE25009
ID AAE25009 standard; protein; 134 AA.

XX AC AAE25009;

XX 30-OCT-2002 (first entry)

XX Human Id2 protein.

XX Paediatric neoplasm; therapy; Id2; neuroblastoma; Wilm's tumour;

XX retinoblastoma; leukaemia; human.

XX Homo sapiens.

XX WO200249502-A2.

XX 27-JUN-2002.

XX 19-DEC-2001; 2001WO-US049163.

XX 21-DEC-2000; 2000US-0257847P.

XX (YESH) UNIV YESHIVA EINSTEIN COLLEGE.

XX Iavarone A, Lasorella A;

XX WPI; 2002-508717/54.

XX N-PSDB; AAD40553.

XX Determining pediatric neoplasm, assessing efficacy of pediatric neoplasm therapy in a subject, or assessing prognosis of pediatric neoplasm in a subject, by assaying for Id2 expression in subject's sample.

XX Disclosure; Fig 13; 79pp; English.

XX The invention relates to a method of determining whether a subject has paediatric neoplasm. The invention also provides methods for assessing the efficacy of therapy to treat a paediatric neoplasm in a subject who has undergone or is undergoing treatment for a paediatric neoplasm; or assessing the prognosis of a subject who has a paediatric neoplasm; by assaying diagnostic sample of the subject for Id2 expression. Methods of the invention are used for treating paediatric neoplasms e.g. Wilm's tumour, neuroblastoma, retinoblastoma, leukaemia and tumours of the central nervous system. The present sequence is human Id2 protein.

XX Sequence 134 AA;

Abp96792 Human COP
Ade61039 Human Pro
Add45104 Human Pro
Ade54569 Human Pro
Aar38910 HEIR-1.. 3
Ade61037 Rat Prote
Ade54567 Rat Prote
Add45102 Rat Prote
Aab19318 Amino aci
Aae37791 Mouse dia
Aab58749 Breast an
Aab41360 Human DIT
Ade61037 Human ORF
Ade54567 Human ORF
Aab19327 Amino aci
Aab57740 Drosophil
Aab19322 Amino aci
Aab28251 Novel hum
Aab58369 Lung canc
Aab63436 Drosophil
Aab53638 Human col

Query Match 97.7%; Score 670.5; DB 5; Length 134;
Best Local Similarity 99.3%; Pred. No. 3.3e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVRKNSLSHSLGSRKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVYS 60
DB 1 MKAFSPVRSVRKNSLSHSLGSRKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVYS 60

QY 61 KWEILOHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTNTDISILSQASE 120
DB 61 KWEILOHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTNTDISILSQASE 119

QY 121 FPELSMNSDKALCG 135
DB 120 FPELSMNSDKALCG 134

RESULT 2
ID AAB19323 standard; protein; 134 AA.
AC AAB19323;
DT 19-FEB-2001 (first entry)
DE Amino acid sequence of an active mimic of human Id2.
KW Id protein; transcription factor; cellular differentiation;
KW cell proliferation; apoptosis; immunogen.
XX Homo sapiens.
OS US6127178-A.
PN 03-OCT-2000.
PD 20-MAR-1998; 98US-00045764.
PF 20-MAR-1998; 98US-00045764.
PR (RECC) UNIV CALIFORNIA.
PA Israel MA, Florio M;
PI WPI; 2000-627985/60.
DR Modulating apoptosis of an isolated target cell having undesirable level
PT of apoptotic induction comprising introducing polynucleotide encoding N-
PT terminal domain of a natural Id protein and confirming the modulation.
XX Claim 11; Col 25-26; 19pp; English.

The present sequence represents an active mimic of Id2. Id proteins
comprise a family of helix-loop-helix (HLH) transcription factors that
are important regulators of cellular differentiation and proliferation.
Id proteins lack a basic DNA binding region and are capable of inhibiting
gene expression. Id gene expression is enhanced in response to mitogenic
stimuli. Id proteins are used to modulate apoptosis of an isolated target
cell having an undesirable level of apoptotic induction. Polypeptides
comprising an N-terminal domain of a natural Id protein are useful as
immunogens, targets in screening assays, bioactive reagents for
modulating cell growth, differentiation and/or function

Query Match 96.4%; Score 661.5; DB 3; Length 134;
Best Local Similarity 97.8%; Pred. No. 3.9e-70;
Matches 132; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVRKNSLSHSLGSRKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVYS 60
DB 1 MKAFSPVRSVRKNSLSHSLGSRKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVYS 60

QY 61 KWEILOHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTNTDISILSQASE 120
DB 61 KWEILOHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTNTDISILSQASE 119

QY 121 FPELSMNSDKALCG 135
DB 120 FPELSMNSDKALCG 134

RESULT 3
ID AAB56969 standard; protein; 176 AA.
AC AAB56969;
DT 13-MAR-2001 (first entry)
DE Human prostate cancer antigen protein sequence SEQ ID NO:1574
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX Homo sapiens.
OS WO200055174-A1.
PN 21-SEP-2000.
PD 08-MAR-2000; 2000WO-US005988.
PF 12-MAR-1999; 99US-0124270P.
PR (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
PI Rosen CA, Ruben SM;
DR WPI; 2000-587513/55.
DR N-PSDB; AAF16172.
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
PS Claim 11; Page 1988; 2338pp; English.

AAF15566 to AAF16505 encode the human prostate cancer associated
proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
The prostate cancer antigens can have neuroprotective, cytostatic,
cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
nephrotropic, antiinfective, gynaecological and antibacterial activities,
and can be used in gene therapy. The prostate cancer antigen
polynucleotides may be used for detection of prostate cancer, chromosome
identification, as chromosome markers, and for numerous other diagnostic
or research purposes. The prostate cancer antigens may be used to treat
disorders such as neural, immune, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
AAF16515 represent sequences used in the exemplification of the present
invention

Query Match 96.4%; Score 661.5; DB 3; Length 176;
Best Local Similarity 97.8%; Pred. No. 5.8e-70;
Matches 132; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVRKNSLSHSLGSRKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVYS 60

Matches	130;	Conservative	2;	Mismatches	2;	Indels	1;	Gaps	1;
QY	1	MKAFSPVRSVRKNSLS	DHSLGIGSR	SKTPVDDP	MSLLYNNM	DCYSKIKELVPS	IPQNKKVS	60	
Db	1	MKAFSPVRSVRKNSLS	DHSLGIGSR	SKTPVDDP	MSLLYNNM	DCYSKIKELVPS	IPQNKKVT	60	
QY	61	KMETLOHLIDYILDQ	LALDSHPTIVSL	HHORPQ	QNSRUTPL	TTLNTD	SILSLQASE	120	
Db	61	KMETLOHLVIDYILD	QALDSHPTIVSL	HHORPQ	QNTSR-TPL	TTLNTD	SILSLQASE	119	
QY	121	FPSELMSNDSKALCG	135						
Db	120	FPSELMSNDSKVLGC	134						
RESULT 6									
ADD47771	ID	ADD47771	standard; protein; 134	AA.					
XX	AC	ADD47771;							
XX	DT	29-JAN-2004	(first entry)						
XX	DE								
XX	DE								
XX	KW	Rat Protein NP_037192,	SEQ ID NO 13467.						
XX	KW	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;							
XX	KW	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.							
XX	OS	Rattus norvegicus.							
XX	XX	WO2003016475-A2.							
XX	XX	27-FEB-2003.							
XX	XX	14-AUG-2002; 2002WO-US025765.							
XX	XX	14-AUG-2001; 2001US-0312147P.							
PR	PR	01-NOV-2001; 2001US-0346382P.							
PR	PR	26-NOV-2001; 2001US-0333347P.							
XX	XX	(GEHO) GEN HOSPITAL CORP.							
PA	PA	(FARB) BAYER AG.							
XX	XX	Woolf C, D'urso D, Befort K, Costigan M;							
PI	PI	WPI: 2003-268312/26.							
DR	DR	GENBANK; NP_037192.							
XX	XX	New composition comprising two or more isolated polypeptides, useful for							
PT	PT	preparing a medicament for treating pain in an animal.							
XX	XX	Claim 1; Page; 1017pp; English.							
XX	XX	The invention discloses a composition comprising two or more isolated rat							
CC	CC	or human polynucleotides or a polynucleotide which represents a fragment,							
CC	CC	derivative or allelic variation of the nucleic acid sequence. Also							
CC	CC	claimed are a vector comprising the novel polynucleotide, a host cell							
CC	CC	comprising the vector, a method for identifying a nucleotide sequence							
CC	CC	which is differentially regulated in an animal subjected to pain and a							
CC	CC	kit to perform the method an array, a method for identifying an agent							
CC	CC	that increases or decreases the expression of the polynucleotide sequence							
CC	CC	that is differentially expressed in neuronal tissue of a first animal							
CC	CC	subjected to pain, a method for identifying a compound which regulates							
CC	CC	the expression of a polynucleotide sequence which is differentially							
CC	CC	expressed in an animal subjected to pain, a method for identifying a							
CC	CC	compound that regulates the activity of one or more of the							
CC	CC	polynucleotides, a method for producing a pharmaceutical composition, a							
CC	CC	method for identifying a compound or small molecule that regulates the							
CC	CC	activity in an animal of one or more of the polypeptides that regulates the							
CC	CC	specification, a method for identifying a compound useful in treating							
CC	CC	pain and a pharmaceutical composition comprising the one or more							
CC	CC	polypeptides or their antibodies. The polynucleotide or the compound that							
CC	CC	modulates its activity is useful for preparing a medicament for treating							

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 134 AA;

Query Match 95.4%; Score 654.5; DB 7; Length 134;
Best Local Similarity 96.3%; Pred. No. 2.6e-69;
Matches 130; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVRKNSLDHSHSIGISRSKTVPDDPMSLLYNNDSCYSKLKELVPSIPONKKVS 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MKAFSPVRSVRKNSLDHSHSIGISRSKTVPDDPMSLLYNNDSCYSKLKELVPSIPONKKVT 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 KMEILQLHLIDYLDLQIALDSHPITVSLHHQRPGQNORSRTTPLTTLTNTDISLSLQASE 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 KMEILQHVIDYILDQLALDSHPITVSLHHQRPGQNQTSR-TPLTTLTNTDISLSLQASE 119
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 FPSELMGNSDKALCG 135
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 120 FPSELMGNSKVLCG 134
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
AAB19317
ID AAB19317 standard; protein; 134 AA.
XX AC AAB19317;
XX DT 19-FEB-2001 (first entry)
XX DE Amino acid sequence of Id2 protein.
XX KW Id protein; transcription factor; cellular differentiation;
XX KW cell proliferation; apoptosis; immunogen.
XX OS Mus sp.
XX PN US6127178-A.
XX PD 03-OCT-2000.
XX PF 20-MAR-1998; 98US-00045764.
XX PR 20-MAR-1998; 98US-00045764.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Israel MA, Florio M;
XX WPI; 2000-627985/60.
XX PT Modulating apoptosis of an isolated target cell having undesirable level
XX PT of apoptotic induction comprising introducing polynucleotide encoding N-
XX PT terminal domain of a natural Id protein and confirming the modulation.
XX PS Claim 11; Col 19-20; 19pp; English.
XX CC The present sequence represents an exemplary Id protein. Id proteins
XX CC comprise a family of helix-loop-helix (HLH) transcription factors that
XX CC are important regulators of cellular differentiation and proliferation.
XX CC Id proteins lack a basic DNA binding region and are capable of inhibiting
XX CC gene expression. Id gene expression is enhanced in response to mitogenic
XX CC stimuli. Id proteins are used to modulate apoptosis of an isolated target
XX CC cell having an undesirable level of apoptotic induction. Polypeptides
XX CC comprising an N-terminal domain of a natural Id protein are useful as
XX CC immunogens, targets in screening assays, bioactive reagents for
XX CC modulating cell growth, differentiation and/or function

```
SQ Sequence 134 AA;
Query Match          94.2%; Score 646.5; DB 3; Length 134;
Best Local Similarity 95.6%; Pred. No. 2.3e-68;
Matches 129; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVRKNSLDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQNKVY 60
DB 1 MKAFSPVRSVRKNSLDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQNKVY 60

QY 61 KWEILQHLIDYILDQIALDSHTPTIVSLHHQPGQNRRTTPTLTNTDILSILQASE 120
DB 61 KWEILQHVLDYILDQIALDSHTPTIVSLHHQPGQNRTR-LTLNTDILSILQASE 119

QY 121 FPELSMNSDKALCG 135
DB 120 FPELSMNSDKVLCG 134

RESULT 8
AAB19326
ID AAB19326 standard; protein; 133 AA.
XX
AC AAB19326;
XX
DT 19-FEB-2001 (first entry)
XX
DE Amino acid sequence of an exemplary Id protein.
XX
KW Id protein; transcription factor; cellular differentiation;
cell proliferation; apoptosis; immunogen.
XX
OS Unidentified.
XX
FN US6127178-A.
XX
PD 03-OCT-2000.
XX
PF 20-MAR-1998; 98US-00045764.
XX
PR 20-MAR-1998; 98US-00045764.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Israel MA, Florio M;
XX
WPI; 2000-627985/60.
XX
Modulating apoptosis of an isolated target cell having undesirable level
of apoptotic induction comprising introducing polynucleotide encoding N-
terminal domain of a natural Id protein and confirming the modulation.
XX
Claim 11; Col 27-28; 19pp; English.
XX
The present sequence represents an exemplary Id protein. Id proteins
comprise a family of helix-loop-helix (HLH) transcription factors that
are important regulators of cellular differentiation and proliferation.
XX
Id proteins lack a basic DNA binding region and are capable of inhibiting
gene expression. Id gene expression is enhanced in response to mitogenic
stimuli. Id proteins are used to modulate apoptosis of an isolated target
cell having an undesirable level of apoptotic induction. Polypeptides
comprising an N-terminal domain of a natural Id protein are useful as
immunogens, targets in screening assays, bioactive reagents for
modulating cell growth, differentiation and/or function
XX
Sequence 133 AA;
Query Match          93.3%; Score 640; DB 3; Length 133;
Best Local Similarity 96.3%; Pred. No. 1.4e-67;
Matches 130; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 1 MKAFSPVRSVRKNSLDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQNKVY 60
DB 1 MKAFSPVRSVRKNSLDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQNKVY 60

QY 61 KWEILQHLIDYILDQIALDSHTPTIVSLHHQPGQNRRTTPTLTNTDILSILQASE 120
DB 61 KWEILQHVLDYILDQIALDSHTPTIVSLHHQPGQNRTR-LTLNTDILSILQASE 119

QY 121 FPELSMNSDKALCG 135
DB 120 FPELSMNSDKVLCG 134

RESULT 9
AAB19325
ID AAB19325 standard; protein; 135 AA.
XX
AC AAB19325;
XX
DT 19-FEB-2001 (first entry)
XX
DE Amino acid sequence of an exemplary Id protein.
XX
KW Id protein; transcription factor; cellular differentiation;
cell proliferation; apoptosis; immunogen.
XX
OS Unidentified.
XX
FN US6127178-A.
XX
PD 03-OCT-2000.
XX
PF 20-MAR-1998; 98US-00045764.
XX
PR 20-MAR-1998; 98US-00045764.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Israel MA, Florio M;
XX
WPI; 2000-627985/60.
XX
Modulating apoptosis of an isolated target cell having undesirable level
of apoptotic induction comprising introducing polynucleotide encoding N-
terminal domain of a natural Id protein and confirming the modulation.
XX
Claim 11; Col 27-28; 19pp; English.
XX
The present sequence represents an exemplary Id protein. Id proteins
comprise a family of helix-loop-helix (HLH) transcription factors that
are important regulators of cellular differentiation and proliferation.
XX
Id proteins lack a basic DNA binding region and are capable of inhibiting
gene expression. Id gene expression is enhanced in response to mitogenic
stimuli. Id proteins are used to modulate apoptosis of an isolated target
cell having an undesirable level of apoptotic induction. Polypeptides
comprising an N-terminal domain of a natural Id protein are useful as
immunogens, targets in screening assays, bioactive reagents for
modulating cell growth, differentiation and/or function
XX
Sequence 135 AA;
Query Match          76.1%; Score 522; DB 3; Length 135;
Best Local Similarity 79.3%; Pred. No. 1.4e-53;
Matches 107; Conservative 11; Mismatches 13; Indels 4; Gaps 2;

QY 1 MKAFSPVRSVRKNS--LSDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQMKK 58
DB 1 MKAFSPVRSVRKNS--LSDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIPQMKK 60

QY 59 VSKWEILQHLIDYILDQIALDSHTPTIVSLHHQPGQNRRTTPTLTNTDILSILQOA 118
DB 61 VSKWEILQHVLDYILDQIALDSHVAITSHHHPRFG--QATPRTPLTTLNTDILSILQ 118

QY 119 SEPTSELSMNSDKAL 133
DB 119 SEPTSELSMNSDKAL 133
```

```

Db      119 PEFPSDLITDSDRTL 133

RESULT 10
AA17266
ID AAY17266 standard; peptide; 59 AA.
XX
XX
AC AAY17266;
XX
XX 09-AUG-1999 (first entry)
XX
XX HLH domain of Id2A.
XX
XX ORF polypeptide; dermatomyositis; polymyositis; inclusion body myositis;
KW sarcoïd myopathy; AZT myopathy; myocardial infarction; ischaemia; ERK-6;
KW extracellular signal-regulated kinase; reperfusion.
XX
XX Homo sapiens.
XX
XX WO9527099-A1.
XX
XX 03-JUN-1999.
XX
XX 23-NOV-1998; 98WO-US025046.
XX
XX 25-NOV-1997; 97US-0066848P.
XX
XX (SUGB-) SUGEN INC.
XX
XX Ullrich A, Giot J;
XX
XX WPI; 1999-357837/30.
XX
XX Nucleic acid encoding ORF polypeptide.
XX
XX Disclosure; Fig 2; 79pp; English.
XX
XX The invention relates to a human ORF polypeptide, a substrate for
CC extracellular signal-regulated kinase, ERK-6. Substances that modulate
CC the activity of the ORF polypeptide can be used to treat diseases
CC selected from dermatomyositis, polymyositis, inclusion body myositis,
CC sarcoïd myopathy, AZT myopathy, myocardial infarction, and ischaemia/
CC reperfusion. The probes and antibodies can be used to detect the presence
CC of ORF in a sample
XX
XX Sequence 59 AA;
Query Match 43.3%; Score 297; DB 2; Length 59;
Best Local Similarity 98.3%; Pred. No. 2.2e-27;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 22 IGRSKTPVDDPMSLLYNNMNDYCKLKLVPSPQNKVKSMELQHLIDYILDQLD 80
DB 1 IGRSKTPVDDPMSLLYNNMNDYCKLKLVPSPQNKVKSMELQHLIDYILDQLD 59

RESULT 11
AAB19320
ID AAB19320 standard; protein; 118 AA.
XX
XX AAB19320;
XX
XX 19-FEB-2001 (first entry)
XX
XX Amino acid sequence of a Xenopus Id protein, designated XidX.
XX
XX Id protein; transcription factor; cellular differentiation;
KW cell proliferation; apoptosis; immunogen.
XX
XX Xenopus sp.
XX
XX US6127178-A.
XX

```

```

PD 03-OCT-2000.
XX
XX 20-MAR-1998; 98US-00045764.
XX
XX 20-MAR-1998; 98US-00045764.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX PI Israel MA, Florio M;
XX
XX WPI; 2000-627985/60.
XX
XX Modulating apoptosis of an isolated target cell having undesirable level
PT of apoptotic induction comprising introducing polynucleotide encoding N-
PT terminal domain of a natural Id protein and confirming the modulation.
XX
XX Claim 11; Col 23-24; 19pp; English.
XX
XX The present sequence represents an exemplary Id protein. Id proteins
CC comprise a family of helix-loop-helix (HLH) transcription factors that
CC are important regulators of cellular differentiation and proliferation.
CC Id proteins lack a basic DNA binding region and are capable of inhibiting
CC gene expression. Id gene expression is enhanced in response to mitogenic
CC stimuli. Id proteins are used to modulate apoptosis of an isolated target
CC cell having an undesirable level of apoptotic induction. Polypeptides
CC comprising an N-terminal domain of a natural Id protein are useful as
CC immunogens, targets in screening assays, bioactive reagents for
CC modulating cell growth, differentiation and/or function
XX
XX Sequence 118 AA;
Query Match 38.8%; Score 266.5; DB 3; Length 118;
Best Local Similarity 46.9%; Pred. No. 2.6e-23;
Matches 67; Conservative 17; Mismatches 24; Indels 35; Gaps 6;

QY 1 MKAFSPVRSVRK-----NSLSHSLGIGRS---KTP-VDDPMSLLYNNMNDYCKLKLVP 51
DB 1 MKAFSPVRSMSYCYQAVCCLSQSLSIARGSHKGMDEPMGLLYDMNGCYCKLKLVP 60
QY 52 STPQNKVKSMELQHLIDYILDQLDLSHTIVSLHHQRQGNQRSTTTLTINTDI 111
DB 61 GIPOQSKLSQVILQHVIDIYIDLQIVL-----GEDQQQS----- 95
QY 112 SILSLQAGEFPSELMSNDSKALC 134
DB 96 SILSLQKSDP-SELATQGDTSVC 117

RESULT 12
AAB19321
ID AAB19321 standard; protein; 128 AA.
XX
XX AAB19321;
XX
XX 12-SEP-2003 (revised)
DT 19-FEB-2001 (first entry)
XX
XX Amino acid sequence of a Zebrafish Id protein, designated Zid.
XX
XX Id protein; transcription factor; cellular differentiation;
KW cell proliferation; apoptosis; immunogen.
XX
XX Danio rerio.
XX
XX US6127178-A.
XX
XX 03-OCT-2000.
PD
XX 20-MAR-1998; 98US-00045764.
XX
XX 20-MAR-1998; 98US-00045764.
XX
XX (REGC ) UNIV CALIFORNIA.
XX

```

```

XX  Israel MA, Florio M;
XX  WPI; 2000-627995/60.
XX
XX  Modulating apoptosis of an isolated target cell having undesirable level
XX  of apoptotic induction comprising introducing polynucleotide encoding N-
XX  terminal domain of a natural Id protein and confirming the modulation.
XX
XX  Claim 11; Col 23-24; 19pp; English.
XX
XX  The present sequence represents an exemplary Id protein. Id proteins
XX  comprise a family of helix-loop-helix (HLH) transcription factors that
XX  are important regulators of cellular differentiation and proliferation.
XX  Id proteins lack a basic DNA binding region and are capable of inhibiting
XX  gene expression. Id gene expression is enhanced in response to inhibiting
XX  stimuli. Id proteins are used to modulate apoptosis of an isolated target
XX  cell having an undesirable level of apoptotic induction. Polypeptides
XX  comprising an N-terminal domain of a natural Id protein are useful as
XX  immunogens, targets in screening assays, bioactive reagents for
XX  modulating cell growth, differentiation and/or function. (Updated on 12-
XX  SEP-2003 to standardise OS field)
XX
XX  Sequence 128 AA;
XX
XX  Query Match          36.8%; Score 252.5; DB 3; Length 128;
XX  Best Local Similarity 41.8%; Pred. No. 1.3e-21;
XX  Matches 61; Conservative 24; Mismatches 30; Indels 31; Gaps 6;
XX
XX  QY 1 MKAFSPVRSVRKNS-----LSDHSLGIGRSKTP-VDDPMSL-LYNNMDCYSKLKEL 49
XX  Db 1 MKVVGFTCALSKSVGGEDVVRCLSDQSLAISKICKIPLLDEQMTWFLQDMNCSYKLKEL 60
XX
XX  QY 50 VPSIPONKVKSMKEILQHLIDYLLDQALDQSHPTIVSLHQRPCNQQRSR-TTPTLTILN 108
XX  Db 61 VFTHTNKKASKMEILQHVLDYIDWLQVELESK-----KNQTSAPRTPLTILN 108
XX
XX  QY 109 TDISILSLQASEFPSELMNSDKALC 134
XX  Db 109 AELASISVENG-----CSDDRINC 127
XX
XX  RESULT 13
XX  ABP41248
XX  ID ABP41248 standard; protein; 229 AA.
XX
XX  AC ABP41248;
XX
XX  DT 23-AUG-2002 (first entry)
XX
XX  DE Human ovarian antigen HOENX16, SEQ ID NO:2380.
XX
XX  KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX  KW ovarian cancer; breast cancer; reproductive system disorder; infertility;
XX  KW pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS;
XX  KW ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX  KW inflammatory condition; immune disorder; blood disorder;
XX  KW cardiovascular disorder; respiratory disorder; neurological disorder;
XX  KW gastrointestinal disorder; urinary system disorder; drug screening;
XX  KW gene therapy; chromosome mapping; forensic analysis;
XX  KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX  KW antiinflammatory; gynaecological; reproductive; chromosome 6p22-21.
XX
XX  OS Homo sapiens.
XX
XX  EN WO2000200677-A1.
XX
XX  PD 03-JAN-2002.
XX
XX  PF 07-JUN-2001; 2001WO-US018569.
XX
XX  FR 07-JUN-2000; 2000US-0209467P.
XX

```

```

PA  (HUMA-) HUMAN GENOME SCI INC.
XX  Birse CE, Rosen CA;
XX  WPI; 2002-147878/19.
XX  DR N-PSDB; ABQ54325.
XX
XX  Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX  useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX  cancer), immune disorders, cardiovascular disorders and neurological
XX  diseases.
XX
XX  Claim 11; SEQ ID NO 2380; 2922pp; English.
XX
XX  The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX  ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX  encompasses polypeptides 90% identical and polynucleotides 95% identical
XX  to the sequences of the invention. The invention additionally relates to
XX  recombinant vectors and host cells comprising human ovarian antigen
XX  polynucleotides, antibodies against human ovarian antigens, and the use
XX  of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX  treating, prognosing or preventing various ovary and/or breast-related
XX  disorders. Such conditions include ovarian cancer and breast cancer, and
XX  metastatic tumours of ovarian or breast origin, reproductive system
XX  disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX  polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX  disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX  shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX  vaginitis), immune disorders (e.g., congenital and acquired
XX  immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX  blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX  respiratory disorders, neurological disorders, gastrointestinal disorders
XX  and urinary system disorders. Ovarian antigen polypeptides and
XX  polynucleotides may also be used in screening for compounds which
XX  modulate ovarian antigen expression or activity. The polynucleotides may
XX  further be used for gene therapy, chromosome mapping, in the
XX  identification of individuals and in forensic analysis, and the
XX  polypeptides may be used as food additives or to prepare antibodies
XX  useful in disease diagnosis, drug targeting and phenotyping. The present
XX  sequence represents a human ovarian antigen of the invention. Note: The
XX  sequence data for this patent did not form part of the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences
XX
XX  Sequence 229 AA;
XX
XX  Query Match          36.3%; Score 249; DB 5; Length 229;
XX  Best Local Similarity 41.8%; Pred. No. 8e-21;
XX  Matches 61; Conservative 17; Mismatches 32; Indels 36; Gaps 4;
XX
XX  QY 1 MKAFSPVRSVRKNSLS-----DHSLGIS-----RSKTPVDDP 32
XX  Db 69 MKAVSPVRPSGRKAPSGCGGELALRCLAEHGHSLGGSAASAAAAARCKAAEAADP 128
XX
XX  QY 33 -MSLLYNNDCYSKLEIVPSIPONKVKSMKEILQHLIDYLLDQALDQSHPTIVSL--- 88
XX  Db 129 ALCLQCDMDCYSKLRRLVPTIPENKVKSVKEILQHVLDYIDWLQVELESK----- 188
XX
XX  QY 89 ---HHQRPGQNQRSTRTPLTNTD 110
XX  Db 189 PAPPHHAGTCPPAPPTPLTALNTD 214
XX
XX  RESULT 14
XX  AAR53710
XX  ID AAR53710 standard; protein; 154 AA.
XX
XX  AC AAR53710;
XX
XX  DT 25-MAR-2003 (revised)
XX  DT 23-JAN-1995 (first entry)
XX
XX  DE Human Id-1H protein.

```

XX myogenic regulatory factors; transcriptional regulatory factors;
 KW diagnosis; therapy; cell proliferation; antisense therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP60627-A1.
 XX
 PD 08-JUN-1994.
 XX
 PF 15-NOV-1993; 93EP-00309097.
 XX
 PR 13-NOV-1992; 92JP-00328391.
 XX
 PA (SUME) SUMITOMO ELECTRIC IND CO.
 XX
 PI Oda K, Nakada S, Hara E, Yamaguchi T, Nakamura T, Oka Y;
 PI Kishimoto T;
 XX
 DR WPI; 1994-177938/22.
 DR N-PSDB; AAQ66082.
 XX
 PT Novel human Id genes - obtd. by probing cDNA library of human TIG-1
 PT fibroblasts with synthetic oligo:nucleotide probes based on mouse Id
 PT gene's conserved regions.
 XX
 PS Claim 1; Page 10-11; 18pp; English.
 XX
 CC AAR63710 shows a novel human Id-1 (Id-H1) protein. Id proteins have
 CC myogenic regulatory factor activity, suppressing the activity of muscle
 CC specific genes. 3 mouse Id genes have been identified and previously only
 CC Id-H2. The 2 human Id genes in the specification are derived from the
 CC same genetic locus on a genome by alternate splicing (see also AAQ66083 -
 CC Id-H1'). The expression of the Id-H1 and Id-H1' genes specifically varies
 CC according to the state of cell proliferation, in particular, the
 CC acquisition of cell aging and permanently proliferating ability. The genes
 CC can be used in diagnostic probes for determining the state of
 CC proliferation and state of differentiation in cells. Antisense molecules
 CC may be utilised as inhibitors of cell proliferation, where the Id
 CC molecules act as transcriptional regulatory factors. (Updated on 25-MAR-
 CC 2003 to correct PN field.)
 XX
 SQ Sequence 154 AA;
 Query Match 35.3%; Score 242.5; DB 2; Length 154;
 Best Local Similarity 45.0%; Pred. No. 2.7e-20;
 Matches 58; Conservative 24; Mismatches 28; Indels 19; Gaps 5;
 QY 15 LSDHSLGISR-----SKTP--VDDPM--SLLYNNDCYSKELVPSIPONKKVSKWEIL 65
 DB 35 LSEQSVASISRCRGARLPALLDEQQVNVLLYDMNGCYSELKELVTLPTLQNRKVSKEIL 94
 QY 66 QHLIDYILDQLALDSHTPTIVSLHQRPGQNRSTRTPLTLNTDITSLQASEPPEL 125
 DB 95 QHVIDYIRDLQLEINSEVGT-----PGRGGLPVRAPLSTLNGEISALTAEACVPA-- 147
 QY 126 MSNDKALC 134
 DB 148 ---DDRILC 153
 RESULT 15
 ID AAU10351 standard; protein; 154 AA.
 XX AAU10351;
 AC
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Human inhibitor of DNA binding-1.
 XX
 KW Human; inhibitor of DNA binding-1; Id-1; cytostatic; antiinflammatory;
 KW immunosuppressive; antisense therapy; antisense oligonucleotide;

KW hyperproliferative disorder; immune disorder; muscular disorder;
 KW vascular disorder; pancreatic disorder; infection; inflammation; tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200183513-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 25-APR-2001; 2001WO-US013209.
 XX
 PR 28-APR-2000; 2000US-00561497.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Baker BF, Bennett CF, Wyatt JR;
 DR WPI; 2002-041477/05.
 DR N-PSDB; AAS16579.
 XX
 PT Novel antisense compound, specifically hybridizing to and inhibiting the
 PT expression of Inhibitor of DNA binding-1, useful for treating
 PT hyperproliferative, immune, muscular, vascular or pancreatic disorder.
 XX
 PS Disclosure; Page 87-88; 105pp; English.
 XX
 CC The invention relates to novel antisense compounds (I) 8-30 nucleobases
 CC in length targeted to a nucleic acid molecule encoding Inhibitor of DNA
 CC binding-1, where (I) specifically hybridises with and inhibits the
 CC expression of Inhibitor of DNA binding-1. Antisense inhibition of human
 CC Inhibitor of DNA binding-1 expression by chimeric phosphorothioate
 CC oligonucleotides having 2'-methoxyethyl (2'-MOE) wings and a deoxy gap
 CC was tested. A series of oligonucleotides were designed to target
 CC different regions of the human Inhibitor of DNA binding-1 RNA. The
 CC compounds were analysed for their effect on human Inhibitor of DNA
 CC binding-1 mRNA levels by quantitative real-time polymerase chain reaction
 CC (PCR). The result showed that the oligonucleotides showed at least 25%
 CC for inhibiting the expression of Inhibitor of DNA binding-1 in cells or
 CC tissues by contacting the cells or tissues with (I). (I) is also useful
 CC for treating a human having a disease or condition associated with
 CC Inhibitor of DNA binding-1 by administering a therapeutically or
 CC prophylactically effective amount of (I), where the disease or condition
 CC is a hyperproliferative disorder, immune disorder, muscular disorder,
 CC vascular disorder or pancreatic disorder. (I) may also be used for
 CC diagnostics, therapeutics, prophylaxis (e.g., to prevent or delay
 CC infection, inflammation or tumour formation), and as research reagents
 CC and kits. (I) may be safely and effectively administered to humans. The
 CC present sequence represents the amino acid sequence of human Inhibitor of
 CC DNA binding-1, the coding sequence of which was used to design the
 CC antisense oligonucleotides of the invention
 XX
 SQ Sequence 154 AA;
 Query Match 35.3%; Score 242.5; DB 5; Length 154;
 Best Local Similarity 45.0%; Pred. No. 2.7e-20;
 Matches 58; Conservative 24; Mismatches 28; Indels 19; Gaps 5;
 QY 15 LSDHSLGISR-----SKTP--VDDPM--SLLYNNDCYSKELVPSIPONKKVSKWEIL 65
 DB 35 LSEQSVASISRCRGARLPALLDEQQVNVLLYDMNGCYSELKELVTLPTLQNRKVSKEIL 94
 QY 66 QHLIDYILDQLALDSHTPTIVSLHQRPGQNRSTRTPLTLNTDITSLQASEPPEL 125
 DB 95 QHVIDYIRDLQLEINSEVGT-----PGRGGLPVRAPLSTLNGEISALTAEACVPA-- 147
 QY 126 MSNDKALC 134
 DB 148 ---DDRILC 153
 Search completed: September 27, 2004, 11:27:24
 Job time : 58 secs

B/G/K

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 11:19:16 ; Search time 39 Seconds

(without alignments)
1092.178 Million cell updates/sec

Title: US-10-025-170-2

Perfect score: 686

Sequence: 1 MKAFSPVRSVRKNSLSDBSL.....LQASEPSELMNSDKALCG 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629.5	91.8	134	13	073933
2	573.5	83.6	376	11	Q7TP14
3	566.5	82.6	168	11	Q9CYW7
4	532	77.6	133	13	Q9PWJ5
5	532	77.6	162	13	Q8AVD1
6	526	76.7	133	13	Q9YGL0
7	522	76.1	135	13	Q42448
8	511	74.5	137	13	Q7S2Q2
9	467	68.1	141	13	Q9W619
10	350.5	51.1	130	13	Q7T164
11	305.5	44.5	131	13	Q7Z2B5
12	305.5	44.5	131	13	Q7ZXF3
13	292	42.6	125	13	Q90X14
14	269.5	39.3	116	13	Q8QFX4
15	268	39.1	127	13	Q9W620
16	266.5	38.8	118	13	Q91399

17	265.5	38.7	118	13	Q91417
18	260	37.9	126	13	Q42447
19	257.5	37.5	128	13	Q42361
20	255.5	37.2	118	13	Q7S228
21	248	36.2	161	11	Q8CH17
22	241.5	35.2	134	13	Q90X12
23	241.5	35.2	140	13	Q90X13
24	212	30.9	129	11	Q922W6
25	209	30.5	129	11	Q8CD45
26	193	28.1	77	13	Q91418
27	166	24.2	118	13	Q90X15
28	152	22.2	36	4	Q14602
29	147.5	21.5	178	5	Q9U905
30	94.5	13.8	108	13	Q9PRF6
31	94.5	13.8	214	13	F70050
32	93	13.6	194	11	O55208
33	93	13.6	196	13	Q90259
34	93	13.6	218	13	Q90764
35	92.5	13.5	195	13	Q90260
36	92.5	13.5	200	13	O13150
37	92	13.4	158	5	Q817T7
38	92	13.4	394	13	Q73823
39	91	13.3	185	5	Q25179
40	91	13.3	376	5	Q9W4X2
41	91	13.3	376	5	O76880
42	90	13.1	309	13	Q919C9
43	89.5	13.0	219	5	Q962A3
44	89.5	13.0	266	5	Q86G55
45	89.5	13.0	348	13	O73687

ALIGNMENTS

RESULT 1

073933 ID 073933 PRELIMINARY; PRT; 134 AA.
AC 073933;
DT 01-AUG-1998 (TREMREL. 07, Created)
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Transcriptional regulator ID2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N. A.
RA Martinsen B.J., Bronner-Fraser M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N. A.
RA Spirin K.S., Kenney M.C., Koeffler H.P.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.
CC EMBL; AF068831; AAC32827.1; -.
CC EMBL; AF049135; AAC05130.1; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00038; HLH_1; 1.
CC PROSITE; PS00888; HLH_2; 1.
SQ SEQUENCE 134 AA; 14798 MW; 98DF7ACEAB2C3562 CRC64;

Query Match 91.8%; Score 629.5; DB 13; Length 134;

Best Local Similarity 91.8%; Pred. No. 1.1e-54;
Matches 124; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MKAFSPVRSVRKNSLSDBSLGIGSRKTPVDPDPSLLYNNMDCYSKLKELVPSIPQNKVS 60

Db 1 MKAFSPVRSVRKNSLSDBSLGIGSRKTPVDPDPSLLYNNMDCYSKLKELVPSIPQNKVS 60

```

Qy 61 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQASE 120
D 61 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQASE 119
Qy 121 FPELSMNSDKALCG 135
D 120 FPELSMNSDKALCG 134

RESULT 2
ID Q7TPI4 PRELIMINARY; PRT; 376 AA.
AC Q7TPI4;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE AC2-300.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Ma H., Wang L., Wang S.F., Han H.P.,
RA Wang G.P., Chai L.Q., Yuan J.Y., Yang K.J., Yan H.M., Chang C.F.,
RA Zhao L.F., Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY321351; AAP86283.1; -.
SQ SEQUENCE 376 AA; 41508 MW; 29E8272D0370D763 CRC64;

Query Match 83.6%; Score 573.5; DB 11; Length 376;
Best Local Similarity 92.7%; Pred. No. 1.3e-48;
Matches 115; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MKAFSPVSRVKNLSLSDHSLGSRKTPVDDPMSLLYNNMDCYKSLKELVPSIPQNKVVS 60
D 53 MKAFSPVSRVKNLSLSDHSLGSRKTPVDDPMSLLYNNMDCYKSLKELVPSIPQNKVVS 112
Qy 61 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQASE 120
D 113 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQGGV 171
Qy 121 FPE 124
D 172 FLNE 175

RESULT 3
ID Q9CYW7 PRELIMINARY; PRT; 168 AA.
AC Q9CYW7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 10, 11 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2810434H03, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

```

```

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AK013239; BAB28736.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS00888; HLH_2; 1.
SQ SEQUENCE 168 AA; 19110 MW; 24998F7B30761CEB CRC64;

Query Match 82.6%; Score 566.5; DB 11; Length 168;
Best Local Similarity 96.6%; Pred. No. 2.5e-48;
Matches 113; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MKAFSPVSRVKNLSLSDHSLGSRKTPVDDPMSLLYNNMDCYKSLKELVPSIPQNKVVS 60
D 14 MKAFSPVSRVKNLSLSDHSLGSRKTPVDDPMSLLYNNMDCYKSLKELVPSIPQNKVVS 73
Qy 61 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQ 117
D 74 KWEILQHLIDYILDQIALDSHPTIVSLHQPONQSRRTPLTTLNTDLSILSQ 129

RESULT 4
ID Q9PWJ5 PRELIMINARY; PRT; 133 AA.
AC Q9PWJ5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Helix-loop-helix type negative regulator Id2.
GN ID2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Regenerating limb blastema;
RA Shimizu-Nishikawa K., Tazawa I., Uchiyama K., Yoshizato K.;
RT "Expression of the helix-loop-helix type negative regulators of
RT differentiation during limb regeneration in urodele and anuran.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AB019520; BAA76634.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS00888; HLH_2; 1.
SQ SEQUENCE 133 AA; 14791 MW; 9CED81FF47D4097 CRC64;

Query Match 77.6%; Score 532; DB 13; Length 133;
Best Local Similarity 80.7%; Pred. No. 5.1e-45;
Matches 109; Conservative 13; Mismatches 9; Indels 4; Gaps 3;

Qy 1 MKAFSPVSRVKNLSLSDHSLGSRKTPVDDPMSLLYNNMDCYKSLKELVPSIPQNKVVS 60

```

```

Db 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKKVS 60
QY 61 KWEILQHLIDYILDQLALDSHPITVSLHH-ORPGONQSRRTPLTTLNTDLSILSLOAS 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 KWEILQHVLDYILDQLTLDSHPSIVSLHHLPRVGGN--TSRTPPLTPLNTDLSILSLOAA 118
QY 120 EFPSELSMNSDKALC 134
Db 119 EFSSE-FTDESKSLC 132

RESULT 5
Q8AVD1 PRELIMINARY; PRT; 162 AA.
ID Q8AVD1
AC Q8AVD1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to inhibitor of DNA binding 2, dominant negative
DE helix-loop-helix protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041527; AAH41527.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
FT NON TER
SQ SEQUENCE 162 AA; 17893 MW; 97CF7FD595C068 CRC64;

Query Match 77.6%; Score 532; DB 13; Length 162;
Best Local Similarity 80.7%; Pred. No. 6.3e-45;
Matches 109; Conservative 13; Mismatches 9; Indels 4; Gaps 3;

QY 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKKVS 60
Db 30 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKKVS 89
QY 61 KWEILQHLIDYILDQLALDSHPITVSLHH-ORPGONQSRRTPLTTLNTDLSILSLOAS 119
Db 90 KWEILQHVLDYILDQLTLDSHPSIVSLHHLPRVGGN--TSRTPPLTPLNTDLSILSLOAA 147
QY 120 EFPSELSMNSDKALC 134
Db 148 EFSSE-FTDESKSLC 161

RESULT 6
Q9YGL0 PRELIMINARY; PRT; 133 AA.
ID Q9YGL0
AC Q9YGL0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE Id2 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Afoula A.B., Reynaud-Deonauth S., Mohun T., Spohr G.B.;

```

```

RT "Localized Xid3 mRNA activation in Xenopus embryos by cytoplasmic
RT Polyadenylation.";
RL Mech. Dev. 85:1-17(1999).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AJ133647; CAB38648.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
SQ SEQUENCE 133 AA; 14805 MW; 9CED81FF470C4187 CRC64;

Query Match 76.7%; Score 526; DB 13; Length 133;
Best Local Similarity 80.0%; Pred. No. 2e-44;
Matches 108; Conservative 13; Mismatches 10; Indels 4; Gaps 3;

QY 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKKVS 60
Db 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKKVS 60
QY 61 KWEILQHLIDYILDQLALDSHPITVSLHH-ORPGONQSRRTPLTTLNTDLSILSLOAS 119
Db 61 KWEILQHVLDYILDQLTLDSHPSIVSLHHLPRVGGN--TSRTPPLTPLNTDLSILSLOAA 118
QY 120 EFPSELSMNSDKALC 134
Db 119 EFSSE-FTDESKSLC 132

RESULT 7
Q42448 PRELIMINARY; PRT; 135 AA.
ID Q42448
AC Q42448;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Id2 protein.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97433269; PubMed=9288909;
RA Rescan P.Y.;
RT "Identification in a fish species of two Id (inhibitor of DNA binding/
RT differentiation)- related helix-loop-helix factors expressed in the
RT slow oxydative muscle fibers.";
RL Eur. J. Biochem. 247:870-876(1997).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; Y08369; CAA69657.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
SQ SEQUENCE 135 AA; 15127 MW; 72C91D6EDA2D4B37 CRC64;

Query Match 76.1%; Score 522; DB 13; Length 135;
Best Local Similarity 79.3%; Pred. No. 5e-44;
Matches 107; Conservative 11; Mismatches 13; Indels 4; Gaps 2;

QY 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKK 58
Db 1 MKAFSPVRSVRKSSLSLTHSGIARSKTPVDDPMSLLYNNMDCYSKELVPSIPONKN 60
QY 59 VSKWEILQHLIDYILDQLALDSHPITVSLHHORPGONQSRRTPLTTLNTDLSILSLOA 118
Db 61 VSKWEILQHVLDYILDQLTLDSHPSIVSLHHLPRVGG--QATPRTPLTTLNTDLSILSLO 118

```

RESULT 9

```

QY 59 VSKWEILOHLDYILDLOIALDSHTTIVSLVHQHRCQONRSRTPPTLTTNTDTSILSLOA 118
Db 57 VSKWEILOHVDIYILDLOIALDONP-LQOOQQSLGQSPPKT--VRSIGADISIISFQP 113
QY 119 SEFFSELSMNSDKAL 133
Db 114 SNPOREINTDDBLIAL 128

```

RESULT 11	
Q7ZZB5	PRELIMINARY; PRT; 131 AA.
Q7ZZB5	
AC Q7ZZB5;	
DT 01-JUN-2003 (trEMBLrel. 24, Created)	
DT 01-JUN-2003 (trEMBLrel. 24, Last sequence update)	
DT 01-OCT-2003 (trEMBLrel. 25, Last annotation update)	
DE Inhibitor of DNA binding 4.	
ID4.	
GN Xenopus laevis (African clawed frog).	
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;	
OC Xenopodinae; Xenopus.	
OX NCBI_TaxID=8355;	
IN [1]	
RN SEQUENCE FROM N.A.	
RP Liu K.J., Harland R.M.;	
RT "Cloning and characterization of Xenopus Id4 reveals differing roles	
RT for Id genes during embryogenesis.";	
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AY279210; AAP34250.1; -.	
DR InterPro; IPR001092; HLH_basic.	
DR Pfam; PF00010; HLH; 1.	
DR SMART; SM00353; HLH; 1.	
DR PROSITE; PS00038; HLH_1; 1.	
DR PROSITE; PS50888; HLH_2; 1.	
DR PROSITE; 131 AA; 14740 MW; 815F3240EB9CC275 CRC64;	
SQ SEQUENCE	

Query Match	44.5%;	Score	305.5;	DB	13;	Length	131;
Best Local Similarity	52.8%;	Pred. No.	1.4e-22;				
Matches	67;	Conservative	17;	Mismatches	22;	Indels	21;
Gaps	3;						
Qy	1	MKAESPVRVKNS-----LSHSLGISRSKTPVDDPMGLILYNNDCYSKLK	47				
Dd	1	MKAVSPRPSRKAAQPSVCGELALHCLSHSLGVARFKMEEEETLCLOYDMNDCYRLK	60				
Qy	48	ELVPSPONKKVKWEILOHLIDYILDQLTALDSHTPTVISLHQRPQGNQRSTPLTLTLL	107				
Dd	61	RLVFTPPNKVKSVKEILLQHVLDLQALDTHP--VLLRQQPPTP-----TPLTDL	112				
Qy	108	NTDISIL	114				
Dd	113	NTDPAAAL	119				

RESULT 12

Q7ZXF3	
ID	PRELIMINARY; PRT; 131 AA.
AC	Q7ZXF3
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Similar to inhibitor of DNA binding 4.
DS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Embryo;
RA	Klein S., Strausberg R.;
RL	Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.

```

DR EMBL; BC045022; AAH45022.1; -.
DR InterPro; IPR001092; HUH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SMO0353; HLH; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS00888; HLH 2; 1.
DR SEQUENCE 131 AA; 14756 MW; 1964B1EB7B822D6B CRC64;

Query Match 44.5%; Score 305.5; DB 13; Length 131;
Best Local Similarity 52.8%; Pred. No. 1.4e-22;
Matches 67; Conservative 17; Mismatches 22; Indels 21; Gaps 3

Qy 1 MKAFSPVVRKNS-----LSCHSLGISRSKTPVDVPMGLLYNMDCYSKUK 47
    ||| |||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 1 MKAVSPVRPQSRKAQVPSVCGELAHCLSEHSLGVARYKYVEEETLCLOYDMNDCYSFLK 60
    ||| |||| : : : ||| : : : ||| : : : ||| : : : ||| : : :

Qy 48 ELVPSIPONKKVSKWEILOHLIDYILDQIALDSHPYIVSLHHQRPQGNQSRRTPLTTL 107
    ||| |||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
Db 61 RLVTPLIPNKKVSKVEILOHVIDYILDQIALDTHP--VLLRQQPPTP-----TELTDL 112
    ||| |||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :

Qy 108 NTDISIL 114
    ||| : ||
Db 113 NTDPAAL 119
    ||| : ||

```

```

RESULT 13
Q90X14 ID Q90X14 PRELIMINARY; PRT; 125 AA.
AC AC Q90X14;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Helix-loop-helix protein.
GN ID3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Kee Y., Bronner-Fraser M.;
RT "The transcriptional regulator Id3 is expressed in cranial sensory
RT placodes during early avian embryonic development.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC EMBL; AY040528; AAK93835.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PFC0010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 125 AA; 133731 MW; 713BA5383A39F313 CRC64;

```

```
Query Match          42.6%; Score 292; DB 13; Length 125;
Best Local Similarity 48.3%; Pred. No. 2.9e-21;
Matches              69; Conservative 21; Mismatches 25; Indels 28; Gaps 5;
```

QY	1	MKAFSVRSVRK-----NSLSDSHSGTSR--KTP-VDDPMSLLYNMDCYKLRELVP	51
Dd	1	MKAISVRSVRSCYEAVCLSEGLAIARSSNNKSALPEEPMNLLVDNMDCYSKLRELVP	60
		: :: :: :: :: :	
QY	52	SIPQNKKVKMEILOHLIDYLQLIALDSHPITIVSLHQRPQGNQRSTPTTLNTDI	111
Dd	61	GIPQGFKLGSVEILQHVIDYIFDLQVLRE-----GAKGRDPSEAT-----	102
		::: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :	
QY	112	STLSLOASFEFPELMNDNKALC	134
Dd	103	-LLSKAAELAGLCCKDERSLC	124
		: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :	

RESULT 14
Q80FX4

```

ID Q8QFX4 PRELIMINARY; PRT; 116 AA.
AC Q8QFX4;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Helix-loop-helix protein.
GN ID3.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=2189534; PubMed=11900982;
RA Dickmeis T., Rastegar S., Lam C.S., Aanstad P., Clark M., Fischer N.,
RA Rosa F., Korzh V., Strahle U.;
RT "Expression of the helix-loop-helix gene id3 in the zebrafish
RT embryo.";
RL Mech. Dev. 113:99-102(2002).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC -!- TRANSCRIPTION FACTORS.
DR EMBL; AY065841; AAL40892.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
SQ SEQUENCE 116 AA; 13115 MW; 7BIACB3B9196AE9 CRC64;

Query Match 39.3%; Score 269.5; DB 13; Length 116;
Best Local Similarity 47.5%; Pred. No. 4.6e-19;
Matches 66; Conservative 15; Mismatches 29; Indels 29; Gaps 4;

QY 1 MKAFSPVRSVRK-----NSLSHSLGIGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIQ 55
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 1 MKAISPVRSVRCYEAVCISQSLAISRCKSPSELSDM--NMNDCYSKIKELVPSIQ 58
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 56 NKKVSKMELIHLIDYLDLQIALDSHTPTIVSLHQRPGQNRSTPTLTNTDISILS 115
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 59 NKSVSQVELIHLQVDFDLQIALENE-----TDTQNTPDFILS 97
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 116 LQASEFPSELMNSDKALC 134
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 98 MKNSEM-SRNFSKEDGAMC 115
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 15
Q9W620 PRELIMINARY; PRT; 127 AA.
AC Q9W620;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Id3 protein.
GN ID3.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
OX NCBI_TaxID=8330;
RN [1]_
RP SEQUENCE FROM N.A.
RX TISSUE-Regenerating limb blastema;
RA Shirizu-Nishikawa K., Tazawa I., Uchiyama K., Yoshizato K.;
RT "Expression of the helix-loop-helix type negative regulators of
RT differentiation during limb regeneration in urodele and anuran.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC -!- TRANSCRIPTION FACTORS.
DR EMBL; AB019515; BAA76632.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.

```

```

SQ SEQUENCE 127 AA; 13883 MW; CCB5162F14251724 CRC64;

Query Match 39.1%; Score 268; DB 13; Length 127;
Best Local Similarity 42.4%; Pred. No. 7.2e-19;
Matches 61; Conservative 26; Mismatches 29; Indels 28; Gaps 5;

QY 1 MKAFSPVRSVRK-----NSLSHSLGIGSRK---TPVDDPMSLLYNNMDCYSKIKELV 51
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 1 MKAISPVRSVRCYEAVCCLSEQSLATARGSGAKGPLEEPMSLLYTMDCYSKIKELV 60
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 52 SIPQKKVSKMELIHLIDYLDLQIALDSHTPTIVSLHQRPGQNRSTPTLTNTDI 111
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 61 GIPOGSQLSQVEILQHVIDYFDLQIVLDEA-----KERPEQ-----GSEA 102
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 112 SILSIQASEP-PSELMNSDKALC 134
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 103 ALLSLKSDFPAPGLTATQERSMC 126
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Search completed: September 27, 2004, 11:28:25
Job time : 42 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 11:17:28 : Search time 10 Seconds
(without alignments)
702.947 Million cell updates/sec

Title: US-10-025-170-2

Perfect score: 686

Sequence: 1 MKAFSPVSRVKNLSLSDHSL.....LQASEFFSELMNSDKALCG 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	661.5	96.4	134	1 ID2 HUMAN	Q02363 homo sapien
2	654.5	95.4	134	1 ID2 MOUSE	P41136 mus musculus
3	654.5	95.4	134	1 ID2 RAT	P41137 rattus norv
4	249	36.3	161	1 ID4 HUMAN	P47928 homo sapien
5	242	35.3	155	1 ID1 HUMAN	P41134 homo sapien
6	241	35.1	161	1 ID4 MOUSE	P41139 mus musculus
7	225.5	32.9	164	1 ID1 RAT	P41135 rattus norv
8	222.5	32.4	168	1 ID1 MOUSE	Q20067 mus musculus
9	217	31.6	119	1 ID3 HUMAN	Q20535 homo sapien
10	214	31.2	119	1 ID3 MOUSE	P41133 mus musculus
11	214	31.2	119	1 ID3 RAT	P41138 rattus norv
12	147.5	21.5	199	1 EMC DROME	P18491 drosophila
13	111	16.2	398	1 TAP DROME	O16867 drosophila
14	92.5	13.5	244	1 NGN1 MOUSE	P70660 mus musculus
15	91.5	13.3	237	1 NGN1 HUMAN	Q32886 homo sapien
16	91.5	13.3	244	1 NGN1 RAT	P70595 rattus norv
17	90	13.1	311	1 SCL CHICK	P24899 gallus gall
18	89.5	13.0	199	1 ASCL1 XENLA	Q62634 xenopus lae
19	88	12.8	233	1 ASCL1 RAT	P19359 rattus norv
20	88	12.8	236	1 ASCL1 HUMAN	P50553 homo sapien
21	86.5	12.6	208	1 NGN1 BRARE	O42606 brachydanio
22	86	12.5	231	1 ASCL1 MOUSE	Q02067 mus musculus
23	85	12.4	133	1 HEN1 HUMAN	Q02575 homo sapien
24	85	12.4	133	1 HEN1 MOUSE	Q02576 mus musculus
25	85	12.4	135	1 HEN2 HUMAN	Q02577 homo sapien
26	85	12.4	135	1 HEN2 MOUSE	Q64221 mus musculus
27	84	12.2	255	1 MYF5 XENLA	P24700 xenopus lae
28	83.5	12.2	192	1 YLB7 CAEEL	P46581 caenorhabdi
29	83.5	12.2	331	1 TAL HUMAN	P17542 homo sapien
30	83	12.1	198	1 AMOS DROME	Q9Y0A7 drosophila
31	82.5	12.0	255	1 MYF5 BOVIN	P17667 bos taurus
32	82	12.0	108	1 TAL2 HUMAN	Q16559 homo sapien
33	82	12.0	197	1 HAN1 XENLA	O73615 xenopus lae

RESULT 1

ID2_HUMAN 34 82 12.0 267 1 LYL1 HUMAN P12980 homo sapien
ID2_MOUSE 35 82 12.0 329 1 TAL_MOUSE P22091 mus musculus
AC Q02363; 36 82 12.0 504 1 PFOM TOBAC Q24164 nicotiana t
DT 01-FEB-1995 (Rel. 31, Last sequence update) 37 80.5 11.7 202 1 HAN1 CHICK Q90691 gallus gall
DT 01-FEB-1995 (Rel. 31, Last sequence update) 38 79.5 11.6 255 1 MYF5_HUMAN P13349 homo sapien
DT 15-MAR-2004 (Rel. 43, Last annotation update) 39 79.5 11.6 255 1 MYF5_MOUSE P24699 mus musculus
DE DNA-binding protein inhibitor ID-2. 40 78.5 11.4 335 1 RSBUT_BAGSU P40399 bacillus su
GN ID2. 41 78 11.4 256 1 MYF5_NOTVT Q91154 notophthalm
OS Homo sapiens (Human). 42 78 11.4 323 1 SY42 ARATH Q95WH4 arabidopsis
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 43 77 11.2 108 1 TAL2 MOUSE Q62282 mus musculus
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo. 44 77 11.2 246 1 MYF5_COTJA P34061 coturnix co
OX NCBI_TaxID=9606; 45 77 11.2 258 1 MYF5_CHICK Q08856 gallus gall

ALIGNMENTS

SEQUENCE FROM N.A.
TISSUE=Lung;
RX MEDLINE=94124570; PubMed=8294468;
RA Hara E., Yamaguchi T., Nojima H., Ide T., Campisi J., Okayama H.,
RA Oda K.;
RT "Id-related genes encoding helix-loop-helix proteins are required for
RT G1 progression and are repressed in senescent human fibroblasts.";
RL J. Biol. Chem. 269:2139-2145(1994).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=92159079; PubMed=1741406;
RP Biggs J., Murphy E.V., Israel M.A.;
RA "A human id-like helix-loop-helix protein expressed during early
RT development.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1512-1516(1992).
RN (3)
RP SEQUENCE FROM N.A.
RX TISSUE=Hippocampus;
MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic
CC DNA-binding domain but are able to form heterodimers with other
CC HLH proteins, thereby inhibiting DNA binding. ID-2 may be an

```

CC      inhibitor of tissue-specific gene expression.
CC      -!- SUBUNIT: Heterodimer with other HLH proteins.
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- TISSUE SPECIFICITY: Highly expressed in early fetal tissues,
CC      including those of the central nervous system.
CC      -!- DEVELOPMENTAL STAGE: Found in most early fetal tissues but not in
CC      the corresponding mature tissues.
CC      -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D13891; BAA02990.1; -.
DR      EMBL; M97796; AAA58681.1; -.
DR      EMBL; BC030639; AAA30639.1; -.
DR      PIR; A40227; A40227.
DR      TRANSFAC; T01212; -.
DR      MIM; 600386; -.
DR      GO; GO:0007275; P:development; TAS.
DR      InterPro; IPR001092; HLH_basic.
DR      Pfam; PF00010; HLH; 1.
DR      SMART; SM00353; HLH; 1.
DR      PROSITE; PS00888; HLH; 1.
KW      Developmental protein; Nuclear protein.
FT      DOMAIN 36 76      HELIX-LOOP-HELIX MOTIF.
FT      CONFLICT 68 76      V -> L (IN REF. 2).
FT      CONFLICT 98 98      A -> R (IN REF. 2).
SQ      SEQUENCE 134 AA; 14917 MW; A0D98B96396EB11E CRC64;

Query Match      96.4%; Score 661.5; DB 1; Length 134;
Best Local Similarity 97.8%; Pred. No. 2.5e-55;
Matches 132; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 MKAFSPVSRVSRKNSLSHSLGSRSKTPVDDPMSLLYNNDCYSKLKELVSPQNKKVS 60
DB      1 MKAFSPVSRVSRKNSLSHSLGSRSKTPVDDPMSLLYNNDCYSKLKELVSPQNKKVS 60
QY      61 KNEILQHLIDYILDQLALDSHPTIVSLHHQPGQNRRTPLTLTNTDISILSQASE 120
DB      61 KNEILQHLIDYILDQLALDSHPTIVSLHHQPGQNRRTPLTLTNTDISILSQASE 119

QY      121 FPELSMNSDKALCG 135
DB      120 FPELSMNSDKALCG 134

RESULT 2
ID2_MOUSE
ID      ID2_MOUSE      STANDARD;      PRT;      134 AA.
AC      P41136; O88604;
DT      01-FEB-1995 (Rel. 31, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      DNA-binding protein inhibitor ID-2.
GN      ID2 OR ID-2 OR IDB2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
SEQUENCE FROM N.A.
RP      MEDLINE=92017841; PubMed=1922066;
RX      Sun X.H., Copeland N.G., Jenkins N.A., Baltimore D.,
RA      "Id proteins Id1 and Id2 selectively inhibit DNA binding by one class
RT      of helix-loop-helix proteins."
RL      Mol. Cell. Biol. 11:5603-5611(1991).
RN      [2]

```

```

RP      SEQUENCE FROM N.A.
RX      MEDLINE=99051333; PubMed=9831657;
RA      Mantani A., Hernandez M.C., Kuo W.-L., Israel M.A.;
RT      "The mouse Id2 and Id4 genes: structural organization and chromosomal
RT      localization."
RL      Gene 222:229-235(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N-3; TISSUE=Mammary gland;
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA      Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC      -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic
CC      DNA-binding domain but are able to form heterodimers with other
CC      HLH proteins, thereby inhibiting DNA binding. ID-2 may be an
CC      inhibitor of tissue-specific gene expression.
CC      -!- SUBUNIT: Heterodimer with other HLH proteins.
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M69293; AAA79771.1; ALT_INIT.
DR      EMBL; AF077860; AAD05214.1; -.
DR      EMBL; BC006921; AAH06921.1; -.
DR      EMBL; BC053699; AAH53699.1; -.
DR      TRANSFAC; T00404; -.
DR      MGD; MGI:96397; Idb2.
DR      InterPro; IPR001092; HLH_basic.
DR      Pfam; PF00010; HLH; 1.
DR      SMART; SM00353; HLH; 1.
DR      PROSITE; PS00888; HLH; 1.
KW      Developmental protein; Nuclear protein.
FT      DOMAIN 36 76      HELIX-LOOP-HELIX MOTIF.
FT      CONFLICT 5 8      SPVR -> RSCE (IN REF. 1).
FT      CONFLICT 102 102      P -> R (IN REF. 1).
SQ      SEQUENCE 134 AA; 14959 MW; A0D98D54F2686CC5 CRC64;

Query Match      95.4%; Score 654.5; DB 1; Length 134;
Best Local Similarity 96.3%; Pred. No. 1.1e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY      1 MKAFSPVSRVSRKNSLSHSLGSRSKTPVDDPMSLLYNNDCYSKLKELVSPQNKKVS 60
DB      1 MKAFSPVSRVSRKNSLSHSLGSRSKTPVDDPMSLLYNNDCYSKLKELVSPQNKKVT 60
QY      61 KNEILQHLIDYILDQLALDSHPTIVSLHHQPGQNRRTPLTLTNTDISILSQASE 120
DB      61 KNEILQHLIDYILDQLALDSHPTIVSLHHQPGQNRRTPLTLTNTDISILSQASE 119

```

```
QY 121 FPELSMNSDKALCG 135
Db 120 FPELSMNSDKVLGC 134

RESULT 3
ID2_RAT ID2_RAT STANDARD; PRT; 134 AA.
AC P41I37;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein inhibitor ID-2.
GN ID2 OR ID-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94197727; PubMed=7908517;
RA Nagata Y., Todokoro K.;
RT "Activation of helix-loop-helix proteins Id1, Id2 and Id3 during
RL Biochem. Biophys. Res. Commun. 199:1355-1362 (1994).
CC -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic
CC DNA-binding domain but are able to form heterodimers with other
CC HLH proteins, thereby inhibiting DNA binding. ID-2 may be an
CC inhibitor of tissue-specific gene expression.
CC -!- SUBUNIT: Heterodimer with other HLH proteins.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10863; BAA01634.1; -.
CC PIR; JC2112; JC2112.
CC TRANSFAC; T01808; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC Developmental protein; Nuclear protein.
CC DOMAIN 36 76 HELIX-LOOP-HELIX MOTIF.
CC SEQUENCE 134 AA; 14989 MW; AD98D54FDEB7B5 CRC64;

Query Match 95.4%; Score 654.5; DB 1; Length 134;
Best Local Similarity 96.3%; Pred. No. 1.1e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKAFSPVRSVKNSLSHSLGISRSKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVVS 60
Db 1 MKAFSPVRSVKNSLSHSLGISRSKTPVDDPMSLLYNNMDCYSKELVPSIPQNKVVS 60

QY 61 KWEILQHLIDYLDLQIALDSDHPTIVSLHQRPGNQPSRTPLTTLNTDLSILSQASE 120
Db 61 KWEILQHLIDYLDLQIALDSDHPTIVSLHQRPGNQPSRTPLTTLNTDLSILSQASE 119

QY 121 FPELSMNSDKALCG 135
Db 120 FPELSMNSDKVLGC 134

RESULT 4
ID4_HUMAN ID4_HUMAN STANDARD; PRT; 161 AA.
```

```

DR EMBL; U28368; AAA73923.1; -.
DR EMBL; Y07958; CAA69255.1; -.
DR EMBL; U16153; AAA82882.1; -.
DR EMBL; AL022726; CAA18779.1; -.
DR EMBL; BC014941; AAI14941.1; -.
DR PIR; G01855; G01855.
DR Genew; HGNC:5363; ID4.
DR MIM; 600581; -.
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW Nuclear protein.
FT DOMAIN 39 48 POLY-ALA.
FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 118 124 POLY-PRO.
FT CONFLICT 10 14 SGRKA -> RPLR (IN REF. 3).
FT CONFLICT 39 40 AA -> Q (IN REF. 3).
FT CONFLICT 77 79 RLV -> WL (IN REF. 3).
SQ SEQUENCE 161 AA; 1662 MW; 5B14847AE7337339 CRC64;

Query Match 36.3%; Score 249; DB 1; Length 161;
Best Local Similarity 41.8%; Pred. No. 1.6e-16;
Matches 61; Conservative 17; Mismatches 32; Indels 36; Gaps 4;

QY 1 MKAFSPVSRVKNLS-----DHSLGIS-----RSKTPVDDP 32
DB 1 MKAFSPVSRGSGGGGGGSGGALRCLAEHGHSLGSAARCAAEAAADDP 60
QY 33 -MSLLYNNDCYSKLKEVPSIPQNKYSKMILOHLIDYIDLQIALDSHTPIVSL--- 88
DB 61 ALCLQDWNDCYSRURRLVPTIPPNKYSKVEILOHVYIDLQALAEHPALLRQPP 120
QY 89 ----HHQRPQGQNRGRTPLTLNTD 110
DB 121 PAPPHHPAGTCAAPPTPLTALNTD 146

RESULT 5
ID_ HUMAN
AC ID1_HUMAN STANDARD; PRT; 155 AA.
DT 01-FEB-1995 (Rel. 31, Created) Q9H0Z5; Q9H109;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-WAR-2004 (Rel. 43, Last annotation update)
DE DNA-binding protein inhibitor ID-1 (ID).
GN ID1 OR ID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS ID-A AND ID-B).
RC TISSUE=lung;
RX MEDLINE=94124570; PubMed=8294468;
RA Hara E., Yamaguchi T., Nojima H., Ide T., Campisi J., Okayama H.,
RA Oda K.;
RT "Id-related genes encoding helix-loop-helix proteins are required for
RT G1 progression and are repressed in senescent human fibroblasts.";
RL J. Biol. Chem. 269:2139-2145(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ID-A).
RC TISSUE=Placenta;
RX MEDLINE=94368847; PubMed=8086456;
RA Deed R.W., Jaslok M., Norton J.D.;
RT "Nucleotide sequence of the cDNA encoding human helix-loop-helix Id-1
RT protein: identification of functionally conserved residues common to
RT Id proteins.";
RL Biochim. Biophys. Acta 1219:160-162(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS ID-A AND ID-B).

```

```

RX MEDLINE=95364622; PubMed=7637581;
RA Zhu W., Dahmen J., Bultone A., Rigole M., Hernandez M.-C., Kuo W.L.,
RA Puellas L., Rubenstein J.L.R., Israel M.A.;
RT "Id gene expression during development and molecular cloning of the
RT human Id-1 gene.";
RL Brain Res. Mol. Brain Res. 30:312-326(1995).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS ID-A AND ID-B).
RC TISSUE=Placenta;
RA Nehlin J.O., Hara E., Kuo W.L., Collins C., Campisi J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehaesialho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Milne S.A., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Sycamore N., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM ID-A).
RC TISSUE=Ovary, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic
CC DNA-binding domain but are able to form heterodimers with other
CC HLH proteins, thereby inhibiting DNA binding.
CC -!- SUBUNIT: Heterodimer with other HLH proteins.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;

```

```

CC Name=ID-A;
CC IsoId=P41134-1; Sequence=Displayed;
CC Name=ID-B;
CC IsoId=P41134-2; Sequence=VSP_002108;
CC -!- DEVELOPMENTAL STAGE: Expression correlates with proliferation in
CC some types of cells.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D13889; BAA02988.1; -
CC EMBL; D13890; BAA02989.1; -
CC EMBL; X77956; CAA54920.1; -
CC EMBL; S78986; AAB35037.1; -
CC EMBL; S78925; AAB35038.1; -
CC EMBL; U57645; AAC13882.1; -
CC EMBL; U57645; AAC13883.1; -
CC EMBL; AL110115; CAC14950.1; -
CC EMBL; AL117381; CAC18793.1; -
CC EMBL; BC000613; AAB00613.1; -
CC EMBL; BC012420; AAB12420.1; -
CC PIR; A49727; A49727.
CC PIR; JC5395; JC5395.
CC PIR; JC5396; JC5396.
CC PIR; S47524; S47524.
CC TRANSFAC; T01801; -.
CC Genew; HGNC:5360; ID1.
CC MIM; 600349; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC Developmental protein; Nuclear protein; Alternative splicing.
CC HELIX-LOOP-HELIX MOTIF.
CC AACVPADRIILCR -> VRSRSDH (in isoform
CC ID-B).
CC /FTid=VSP_002108.
CC -----
CC CONFLICT 16 16 S -> T (IN REF. 1).
CC CONFLICT 46 46 A -> R (IN REF. 1 AND 2).
CC CONFLICT 48 48 MISSING (IN REF. 1 AND 2).
CC SEQUENCE 155 AA; 16132 MW; 480287384B667161 CRC64;
CC -----
CC Query Match 35.3%; Score 242; DB 1; Length 155;
CC Best Local Similarity 44.6%; Pred. No. 6.8e-16;
CC Matches 58; Conservative 24; Mismatches 28; Indels 20; Gaps 5;
CC -----
Qy 15 LSDSLGISR-----SKTP--VDDPM--SLLYNNDCYKELVPSIPONKKVSKMEI 64
Dy 35 LSEQSVAISRCAGGAGALPALDEQVNVLLYDMNGCYSLKELVPLTPQNKVSKVEI 94
Qy 65 LQHLIDYLDLQIALDSHTVLSLHQRPQGNQRSTPLTLTNTDLSILQSEFPSE 124
Dy 95 LQHVLDYRDLQLEINSESEVGT-----PGRGLFVRAPLSTLNGEISALTAEACVPA- 148
Qy 125 LMSNDSKALC 134
Dy 149 ---DDRILC 154
CC -----
RESULT 6
ID4_MOUSE
AC P41139; STANDARD; PRT; 161 AA.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein inhibitor ID-4.

```

```

GN ID4 OR ID-4 OR IDB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/c; TISSUE=Bone marrow;
RX MEDLINE=94188125; PubMed=8139914;
RT Riechmann V., van Cruuchten I., Sablitzky F.;
RT "The expression pattern of Id4, a novel dominant negative helix-loop-
RT helix protein, is distinct from Id1, Id2 and Id3.";
RL Nucleic Acids Res. 22:749-755(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99057583; PubMed=9838043;
RA van Cruuchten I., Cinato E., Fox M., King E.R., Newton J.S.,
RA Riechmann V., Sablitzky F.;
RT "Structure, chromosomal localisation and expression of the murine
RT dominant negative helix-loop-helix Id4 gene.";
RL Biochim. Biophys. Acta 1443:55-64(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99051333; PubMed=9831657;
RA Mantani A., Hernandez M.C., Kuo W.-L., Israel M.A.;
RT "The mouse Id2 and Id4 genes: structural organization and chromosomal
RT localization.";
RL Gene 222:229-235(1998).
CC -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic
CC DNA-binding domain but are able to form heterodimers with other
CC HLH proteins, thereby inhibiting DNA binding.
CC -!- SUBUNIT: Heterodimer with other HLH proteins.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75018; CAA52926.1; -
CC EMBL; AJ001972; CAA05120.1; -
CC EMBL; AF077859; AAD05213.1; -
CC PIR; S43260; S43260.
CC TRANSFAC; T01658; -.
CC MGD; MGI:99414; Idb4.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC Nuclear protein.
CC DOMAIN 39 48 POLY-ALA.
CC DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF.
CC DOMAIN 118 124 POLY-PRO.
CC SEQUENCE 161 AA; 16596 MW; 2DCFF47AF7EE7EED CRC64;
CC -----
CC Query Match 35.1%; Score 241; DB 1; Length 161;
CC Best Local Similarity 42.2%; Pred. No. 8.8e-16;
CC Matches 62; Conservative 17; Mismatches 30; Indels 38; Gaps 5;
CC -----
Qy 1 MKAFSPVRSVRKNSLS-----DHSLGIS-----RSKTPVDDP 32
Dy 1 MKAVSPVRPSGRKAPSKGSGGELALRCLAEHSHSLGGSAAAAAARCKAAEADEP 60
Qy 33 -MSLLYNNDCYKELVPSIPONKKVSKMEILQHLIDYLDLQIALDSHTV----- 86
Dy 61 ALCLQCDMNDYSLRLVPTIPPNKKVSKVEILQHVLDYLDLQIALDTHPALLRQPPP 120
Qy 87 ---SLHHQRPQGNQRSTPLTLTNTD 110
Dy 111

```

Db 121 PAPPLHPAGACFVAPPR-TPLTALNTD 146

RESULT 7

ID1_RAT STANDARD; PRT; 164 AA.

AC P41135;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE DNA-binding protein inhibitor ID-1.

GN ID1 OR ID-1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Heart;

RX MEDLINE=94148974; PubMed=8106493;

RA Springhorn J.P., Singh K., Kelly R.A., Smith T.W.;

RT "Posttranscriptional regulation of Id1 activity in cardiac muscle.

RT Alternative splicing of novel Id1 transcript permits

RT homodimerization.";

RL J. Biol. Chem. 269:5132-5136(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Heart;

RX MEDLINE=92332551; PubMed=1378442;

RA Springhorn J.P., Ellingsen O., Berger H.J., Kelly R.A., Smith T.W.;

RT "Transcriptional regulation in cardiac muscle. Coordinate expression

RT of Id with a neonatal phenotype during development and following a

RT hypertrophic stimulus in adult rat ventricular myocytes in vitro.";

RL J. Biol. Chem. 267:14360-14365(1992).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=94197727; PubMed=7908517;

RA Nagata Y., Todokoro K.;

RT "Activation of helix-loop-helix proteins Id1, Id2 and Id3 during

RT neural differentiation.";

RL Biochem. Biophys. Res. Commun. 199:1355-1362(1994).

CC -!- FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic

CC DNA-binding domain but are able to form heterodimers with other

CC HLH proteins, thereby inhibiting DNA binding.

CC -!- SUBUNIT: Heterodimer with other HLH proteins. The short form

CC (ID1.25) can form homodimers.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=Long;

CC IsoId=P41135-1; Sequence=Displayed;

CC Name=Short; Synonyms=ID1.25;

CC IsoId=P41135-2; Sequence=VSP_002110;

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; L23148; AAA20403.1; -

CC EMBL; D10862; BAA01633.1; -

CC EMBL; M86708; AAA41090.1; -

CC PIR; A53334; A53334.

CC PIR; JC2111; JC2111.

CC TRANSFAC; T01606; -

CC InterPro; IPR001092; HLH_basic.

CC Pfam; PF00010; HLH; 1.

CC SMART; SM00353; HLH; 1.

DR PROSITE; PS50888; HLH; 1.

KW Developmental protein; Nuclear protein; Alternative splicing.

FT DOMAIN 59 99 HELIX-LOOP-HELIX MOTIF

FT VARSPLIC 136 164 VRSESYIILLWETKATGGGPPYFSGA -> AACVPADD

FT RILCR (in isoform Short).

FT /FTID=VSP_002110.

FT CONFLICT 113 113 G -> A (IN REF. 3).

SQ SEQUENCE 164 AA; 17397 MW; C782DDF4C4892D6D CRC64;

Query Match 32.9%; Score 225.5; DB 1; Length 164;

Best Local Similarity 47.9%; Pred. No. 2.6e-14;

Matches 56; Conservative 20; Mismatches 26; Indels 15; Gaps 5;

QY 15 LSDSHSLGTSR---SKTP--VDDPM--SLYNNDCYSKELVPSIPONKVKSKMELIQH 67

DB 31 LSEQVALSRCAGRALPALDEQQNVLLYDMNGCYSKELVFTLPQNRKVKVLEIQH 90

QY 68 LIDYLDLQIALDSHFTIVSLHQRPGQNRSTTPTLTNTDISILQASEFPSE 124

DB 91 VIDYIRDLQELNSESEVAT----AGRGLPVRAPLSTLNGEISAL---AAEVRSE 139

RESULT 8

ID1_MOUSE STANDARD; PRT; 168 AA.

AC P20067; Q61101; Q9D897;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE 10-OCT-2003 (Rel. 42, Last annotation update)

DE DNA-binding protein inhibitor ID-1 (ID).

GN ID1 OR ID-1 OR IDB1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM SHORT).

RX MEDLINE=9019896; PubMed=2156629;

RA Benazzar R., Davis R.L., Lockshon D., Turner D.L., Weintraub H.;

RT "The protein Id: a negative regulator of helix-loop-helix DNA binding

RT proteins.";

RL Cell 61:49-59(1990).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RC STRAIN=Swiss albino; TISSUE=Brain;

RX MEDLINE=96328261; PubMed=8765747;

RA Hernandez M.C., Andres-Barquín P.J., Israel M.A.;

RT "Molecular cloning of the cDNA encoding a helix-loop-helix protein,

RT mouse Id1B: tissue-specific expression of ID1A and ID1B genes.";

RL Biochim. Biophys. Acta 1308:28-30(1996).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM LONG).

RC STRAIN=C57BL/6J; TISSUE=Small intestine;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Hoffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

DR EMBL; X66924; CAA47360.1; ALT INIT.
DR EMBL; A17546; CAA01342.1; ALT_INIT.
DR EMBL; A17548; CAA01343.1; -.
DR EMBL; AL021154; CAA15950.1; -.
DR EMBL; BC003107; AAH03107.1; -.
DR PIR; I37092; S28529.
DR TRANSFAC; T01809; -.
DR Genew; HGNC:5362; ID3.
DR MIM; 600277; -.
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR GO; GO:007275; P:development; TAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW Nuclear protein; Polymorphism.
FT DOMAIN 41 81 HELIX-LOOP-HELIX MOTIF.
FT VARIANT 105 105 A -> T (in dbSNP:11574).
FT VARIANT /FTID=VAR 016122.
SQ SEQUENCE 119 AA; 12969 MW; 7FC38B56B4DCFFBD CRC64;

Query Match 31.6%; Score 217; DB 1; Length 119;
Best Local Similarity 39.3%; Pred. No. 1.1e-13;
Matches 55; Conservative 24; Mismatches 33; Indels 28; Gaps 6;
QY 1 MKAFSPVRSVRKN--SLSDHSLGISK--TPVDDPMSLLYNNMDCYSKLKELVPSIPQ 55
Db 1 MKALSPVRCGYEAVCCLSERSLATARGKGPAAEPLSLDDMMHCHYSRURELVGVPR 60
QY 56 NKKVSKMEILQHLIDYLDLQIALDSHTPTIVSLHQRPGQNRSTTTLTTLTNTDISILS 115
Db 61 GTQLSQVELQRVIDYILDQLQVL--AEP--PGP-----PDGPHLP 98

QY 116 LQASEFPSEL-MSNDSKALC 134
Db 99 IQTAEIAPVLSNDRSFC 118

RESULT 10
ID3_MOUSE ID3_MOUSE STANDARD; PRT; 119 AA.
AC P41133;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein inhibitor ID-3 (ID-like protein inhibitor HLH 462).
DE ID3 OR ID-3 OR IDB3 OR HLH462.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91156698; PubMed=2000388;
RA Christy B.A., Sanders L.K., Lau L.F., Copeland N.G., Jenkins N.A., Nathans D.;
RT "An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene."
RL Proc. Natl. Acad. Sci. U.S.A. 88:1815-1819(1991).
CC FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic DNA-binding domain but are able to form heterodimers with other HLH proteins, thereby inhibiting DNA binding. ID-3 inhibits the binding of E2A-containing protein complexes to muscle creatine kinase E-box enhancer. May inhibit other transcription factors.
CC SUBUNIT: Homodimer, and heterodimer with other HLH proteins.
CC SUBCELLULAR LOCATION: Nuclear.
CC INDUCTION: By a variety of mitogenic agents in serum starved cells.
CC SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M60523; AAA37818.1; -.
DR PIR; A39114; A39114.
DR TRANSFAC; T00367; -.
DR MGD; MGI:96398; Idb3.
DR GO; GO:001904; F:protein domain specific binding; IDA.
DR GO; GO:0000122; P:negative regulation of transcription from P...; IDA.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW Nuclear protein.
FT DOMAIN 41 81 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 119 AA; 13089 MW; 49F3B841D2D2BF15 CRC64;

Query Match 31.2%; Score 214; DB 1; Length 119;
Best Local Similarity 40.0%; Pred. No. 2.1e-13;
Matches 56; Conservative 24; Mismatches 32; Indels 28; Gaps 7;
QY 1 MKAFSPVRSVRKN--SLSDHSLGIS--RSKTP--VDDPMSLLYNNMDCYSKLKELVPSIPQ 55
Db 1 MKALSPVRCGYEAVCCLSERSLATARGKSPSTEPLSLDDMMHCHYSRURELVGVPR 60
QY 56 NKKVSKMEILQHLIDYLDLQIALDSHTPTIVSLHQRPGQNRSTTTLTTLTNTDISILS 115
Db 61 GTQLSQVELQRVIDYILDQLQVL--AEP--PGP-----PDGPHLP 98

QY 116 LQASEFPSEL-MSNDSKALC 134
Db 99 IQTAEIAPVLSNDRSFC 118

RESULT 11
ID3_RAT ID3_RAT STANDARD; PRT; 119 AA.
AC P41138;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein inhibitor ID-3.
DE ID3 OR ID-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94197727; PubMed=7908517;
RA Nagata Y., Todokoro K.;
RT "Activation of helix-loop-helix proteins Id1, Id2 and Id3 during neural differentiation."
RL Biochem. Biophys. Res. Commun. 199:1355-1362(1994).
CC FUNCTION: ID (inhibitor of DNA binding) HLH proteins lack a basic DNA-binding domain but are able to form heterodimers with other HLH proteins, thereby inhibiting DNA binding. ID-3 inhibits the binding of E2A-containing protein complexes to muscle creatine kinase E-box enhancer. May inhibit other transcription factors.
CC SUBUNIT: Homodimer, and heterodimer with other HLH proteins.
CC SUBCELLULAR LOCATION: Nuclear.
CC SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195(2000).
[4]
REVIEWS
MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Scapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.B.;
"Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review";
Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
SEQUENCE FROM N.A.
RA STRAIN=Berkeley; TISSUE=Embryo;
MEDLINE=22426066; PubMed=12537569;
RA Scapleton M., Carlson J.W., Brokstein P., Yu C., Champe M., Rubin G.A., Guarini H., Krommiller B., Pacleb J.M., Park S., Wan K.H., George R.A., Celisner S.E.;
"A *Drosophila* full-length cDNA resource";
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RA -1- FUNCTION: Participates in sensory organ patterning by antagonizing the neurogenic activity of the Achaete-scute complex (AS-C). It lacks a basic DNA-binding domain but is able to form heterodimers with other HLH proteins, thereby inhibiting DNA binding. May sequester proneural proteins in complexes inefficient for DNA interaction. EMC also affects vein differentiation. Inhibits the activity of AS-C proteins by forming a non-DNA binding heterodimer.
RA -1- SUBUNIT: Heterodimer with other HLH proteins.
RA -1- SUBCELLULAR LOCATION: Nuclear.
RA -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; M31900; AAA28510.1; -
DR EMBL; M31901; AAA28510.1; JOINED.
DR EMBL; M31902; AAA28509.1; -
EMBL; M32637; AAA28511.1; -
DR EMBL; M32636; AAA28511.1; JOINED.

```

DR EMBL; AE003469; AAF47413.2; -.
DR EMBL; AY069405; AAL39550.1; -.
DR PIR; A34689; A34689.
DR TRANSFAC; T00274; -.
DR FlyBase; FBgn0000575; emc.
DR GO; GO:0007422; P:peripheral nervous system development; IMP.
DR GO; GO:0007461; P:restriction of R8 fate; NAS.
DR GO; GO:0007530; P:sex determination; IGI.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Phosphorylation; Nuclear protein; Repressor; Transcription regulation.
FT DOMAIN 36 76
FT DOMAIN 102 122
FT DOMAIN 161 181
FT MOD RES 106 106
FT CONFLICT 123 123
SQ SEQUENCE 199 AA; 21978 MW; 08C1683352B26F0E CRC64;

Query Match 21.5%; Score 147.5; DB 1; Length 199;
Best Local Similarity 35.4%; Pred. No. 6.7e-07;
Matches 34; Conservative 19; Mismatches 30; Indels 13; Gaps 2;

QY 44 SKLKLVPSPONKVKSMELIQLHLDYLDLQIALDGH-----TIVSLHHQ 91
DB 44 SKLKLVPSPONKVKSMELIQLHLDYLDLQIALDGH-----TIVSLHHQ 103
QY 92 RFGQNRGRTEPLTLNTDTSILSLQASEFFSELM 127
DB 104 EDSMEDADAEAEVDPDILAQLRNA-EQPAKVSS 138

RESULT 13
TAP_DROME STANDARD; PRT; 398 AA.
AC O16867; P91640; Q9VVF9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basic helix-loop-helix neural transcription factor TAP (Protein
DE biparous).
GN TAP OR BPS OR CG7659.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97115720; PubMed=8954743;
RA Bush A.B., Hiromi Y.H., Cole M.D.;
RT "Biparous: a novel bHLH gene expressed in neuronal and glial
RL Dev. Biol. 180:759-772(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Oregon-R.
RC MEDLINE=97354289; PubMed=9210583;
RA Gautier P., Ledent V., Massar M., Dambly-Chaudiere C., Ghysen A.;
RT "tap, a Drosophila bHLH gene expressed in chemosensory organs.";
RL Gene 191:15-21(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards M., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

```

```

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Bhokstein P., Bottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP CHARACTERIZATION.
RC STRAIN=Oregon-R;
RX MEDLINE=98211586; PubMed=9551861;
RA Ledent V., Gaillard F., Gautier P., Ghysen A., Dambly-Chaudiere C.;
RT "Expression and function of tap in the gustatory and olfactory organs
RL of Drosophila.";
RL Int. J. Dev. Biol. 42:163-170(1998).
CC -!- FUNCTION: May play a role in the specification of the sugar-
CC sensitive adult gustatory neuron and affect the response to sugar
CC and salt. Regulated by POXN.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC -----
DR EMBL; AF022883; AAC80572.1; -.
DR EMBL; X95845; CAA65103.1; -.
DR EMBL; AE003524; AAF49352.1; -.
DR FlyBase; FBgn0015550; tap.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Neurogenesis;
KW Developmental protein; Differentiation.
FT DNA BIND 155 166
FT DOMAIN 167 207
FT DOMAIN 18 23
FT DOMAIN 308 312
SQ SEQUENCE 398 AA; 44850 MW; 54FF558483B18258 CRC64;

Query Match 16.2%; Score 111; DB 1; Length 398;
Best Local Similarity 31.1%; Pred. No. 0.0041;
Matches 28; Conservative 20; Mismatches 30; Indels 12; Gaps 1;

```


Search completed: September 27, 2004, 11:27:38
Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 11:21:21 ; Search time 16 Seconds

(without alignments)
811.616 Million cell updates/sec

Title: US-10-025-170-2

Perfect score: 686

Sequence: 1 MKAFSPVRSVKNSLSHDHSL.....LQASEFFPELMSNDKALCG 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	661.5	96.4	134	2 A40227	transcription repr
2	654.5	95.4	134	2 J02212	helix-loop-helix p
3	640	93.3	133	2 J02007	differentiation in
4	621.5	90.6	186	2 A41689	DNA-binding protei
5	266.5	38.8	118	2 I51278	negative regulator
6	265.5	38.7	118	2 I51316	negative regulator
7	249	36.3	161	2 G01855	helix-loop-helix p
8	242.5	35.3	154	2 A49727	DNA-binding protei
9	242.5	35.3	154	2 S47524	gene id1 protein -
10	242	35.3	155	2 J05395	helix-loop-helix p
11	242	35.3	161	2 J03006	Id4 protein - huma
12	241	35.1	161	2 S43260	helix-loop-helix p
13	239.5	34.9	148	2 J02111	helix-loop-helix p
14	235.5	34.3	176	2 A34690	DNA-binding protei
15	227	33.1	148	2 B49727	DNA-binding protei
16	226.5	33.0	149	2 J05396	helix-loop-helix p
17	225.5	32.9	164	2 A53334	helix-loop-helix p
18	222.5	32.4	168	2 S72171	helix-loop-helix p
19	217	31.6	119	2 S28529	helix-loop-helix p
20	214	31.2	119	2 J02113	helix-loop-helix p
21	214	31.2	119	2 A39114	helix-loop-helix p
22	201	29.3	160	2 S71405	helix-loop-helix p
23	193	28.1	77	2 I51318	helix-loop-helix p
24	152	22.2	36	2 J02006	bHLH transcription
25	147.5	21.5	199	2 A34688	extramacrochaetae
26	146.5	21.4	199	2 A34689	extramacrochaetae
27	93	13.6	196	2 I50507	achaete-scute homo
28	92.5	13.5	195	2 I50508	achaete-scute homo
29	90	13.1	311	2 S20085	transcription fact

30	89.5	13.0	199	2 A56548	pro-neural achaete
31	89	13.0	219	2 I51382	achaete-scute homo
32	88	12.8	233	2 S11563	probable MASH-2 pr
33	88	12.8	238	2 A48279	achaete scute prot
34	86	12.5	231	2 S28186	achaete-scute locu
35	85	12.4	133	2 A45075	transcription fact
36	85	12.4	133	2 A41788	basic helix-loop-h
37	85	12.4	135	2 B45075	transcription fact
38	85	12.4	135	2 A49005	basic domain helix
39	84	12.2	255	2 S16151	myogenic factor Xm
40	83.5	12.2	89	2 T29995	hypothetical prote
41	83.5	12.2	192	2 T15764	hypothetical prote
42	83.5	12.2	331	2 A36358	T-cell acute lymph
43	82.5	12.0	255	2 JN0624	Myogenic factor -
44	82	12.0	111	2 A41629	TAL2 protein - hu
45	82	12.0	267	2 A30988	lyl-1 protein - hu

ALIGNMENTS

RESULT 1

A40227

transcription repressor Id-2 - human

N:Alternate names: inhibitor of DNA binding (Id) 2

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000

C:Accession: A40227; C49727

R:Biggs, J.; Murphy, E.V.; Israel, M.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 1512-1516, 1992

A:Title: A human Id-like helix-loop-helix protein expressed during early development.

A:Reference number: A40227; MUID:92159079; PMID:1741406

A:Accession: A40227

A:Molecule type: DNA

A:Residues: 1-134 <BIG>

A>Note: sequence extracted from NCBI backbone (NCBIN:82812, NCBI:82814)

R:Hara, E.; Yanaguchi, T.; Nojima, H.; Ide, T.; Campisi, J.; Okayama, H.; Oda, K.

J. Biol. Chem. 269, 2139-2145, 1994

A:Title: Id-related genes encoding helix-loop-helix proteins are required for G-1 progr

A:Reference number: A49727; MUID:94124570; PMID:8294468

A:Accession: C49727

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-134 <HAR>

A:Cross-references: GB:DL3891; NID:G464183; PIDN:BAA02990.1; PID:G471126

C:Comment: This helix-loop-helix (HLH) protein lacks a basic DNA-binding region and het

C:Genetics:

A:Gene: GDB:ID2

A:Cross-references: GDB:434746; OMTM:600386

A:Map position: 20q11-20q11

C:Superfamily: transcription repressor Id-2

C:Keywords: cell cycle control; transcription factor

Query Match 96.4%; Score 661.5; DB 2; Length 134;
Best Local Similarity 97.8%; Pred. No. 1.2e-56;
Matches 132; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MKAFSPVRSVKNSLSHDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIFQNKVVS 60

Db 1 MKAFSPVRSVKNSLSHDHSLGISRKTVPDDPMSLLYNNMDCYSKLKELVPSIFQNKVVS 60

Qy 61 KMEILQHLIDYLDLQIALDSHPTIVSLHQRPGQNRSTPTLTNTDTSILSLOASE 120

Db 61 KMEILQHLIDYLDLQIALDSHPTIVSLHQRPGQNRSTPTLTNTDTSILSLOASE 119

Qy 121 FPELMSNDKALCG 135

Db 120 FPELMSNDKALCG 134

RESULT 2

JC2112

helix-loop-helix protein, Id2 - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 21-Jul-2000
C:Accession: JC2112; JC2530
R:Nagata, Y.; Todokoro, K.
Biochem. Biophys. Res. Commun. 199, 1355-1362, 1994
A:Title: Activation of helix-loop-helix proteins Id1, Id2 and Id3 during neural differentiation
A:Reference number: JC2111; MUID:94197727; PMID:7908517
A:Accession: JC2112
A:Molecule type: mRNA
A:Residues: 1-134 <NAG>
A:Cross-references: GB:D10863; NID:9434790; PIDN:BAA01634.1; PID:g434791
R:Nagata, Y.; Shoji, W.; Obinata, M.; Todokoro, K.
Biochem. Biophys. Res. Commun. 207, 916-926, 1995
A:Title: Phosphorylation of helix-loop-helix proteins Id1, Id2 and Id3.
A:Reference number: JC2529; MUID:95169152; PMID:7864897
A:Accession: JC2530
A:Molecule type: DNA
A:Residues: 1-134 <NA2>
C:Comment: This protein functions as an inhibitor of cell differentiation and is a nuclear protein
C:Comment: This helix-loop-helix protein forms a heterodimer with ubiquitous and/or tissue-specific proteins
C:Superfamily: transcription repressor Id-2
C:Keywords: differentiation; DNA binding; phosphoprotein; transcription factor
F:35-76/Region: helix-turn-helix motif
F:5,14,25/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
F:5/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted
F:9/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:14/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:14/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
F:16/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
F:27,60/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted
F:27,60/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predicted
F:60/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predicted
F:98/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 95.4%; Score 654.5; DB 2; Length 134;
Best Local Similarity 96.3%; Pred. No. 5.8e-56;
Matches 130; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 60
DB 1 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 60

QY 61 KWEILQHLIDYILDQIALDSHPTIVSLHHQRPQGNQSRRTPLTNTDILSLQASE 120
DB 61 KWEILQHVYILDYILDQIALDSHPTIVSLHHQRPQGNQTSR-TPLTNTDILSLQASE 119

QY 121 FPELSMNSDKALCG 135
DB 120 FPELSMNSDKVLCG 134

RESULT 3
JC2007
differentiation inhibitor protein Id2A - human
C:Species: Homo sapiens (man)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C:Accession: JC2007
R:Kurabayashi, M.; Jeyaseelan, R.; Kedes, L.
Gene 133, 305-306, 1993
A:Title: Two distinct cDNA sequences encoding the human helix-loop-helix protein Id2.
A:Reference number: JC2006; MUID:94040830; PMID:8224921
A:Accession: JC2007
A:Molecule type: mRNA
A:Residues: 1-133 <KOR>
A:Experimental source: heart
C:Comment: This protein functions as a dominant negative regulatory factor.
C:Superfamily: transcription repressor Id-2
C:Keywords: differentiation

Query Match 93.3%; Score 640; DB 2; Length 133;
Best Local Similarity 96.3%; Pred. No. 1.4e-54;
Matches 130; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 1 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 60
DB 1 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 60

QY 61 KWEILQHLIDYILDQIALDSHPTIVSLHHQRPQGNQSRRTPLTNTDILSLQASE 120
DB 61 KWEILQHVYILDYILDQIALDSHPTIVSLHHQRPQGNQASR-TPL-TLNTDILSLQASE 118

QY 121 FPELSMNSDKALCG 135
DB 119 FPELSMNSDKALCG 133

RESULT 4
A1689
DNA-binding protein inhibitor Id-2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
C:Accession: A41689
R:Sun, X.H.; Copeland, N.G.; Jenkins, N.A.; Baltimore, D.
Mol. Cell. Biol. 11, 5603-5611, 1991
A:Title: Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins
A:Reference number: A41689; MUID:92017841; PMID:1922066
A:Accession: A41689
A:Status: preliminary
A:Molecule type: mRNA; DNA
A:Residues: 1-186 <SUN>
A:Cross-references: GB:M69293
C:Superfamily: transcription repressor Id-2

Query Match 90.6%; Score 621.5; DB 2; Length 186;
Best Local Similarity 92.6%; Pred. No. 1.3e-52;
Matches 125; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 60
DB 53 MKAFSPVSRKNSLSHDHSLGSRKTPVDDPMSLLYNNDCYSKELVPSIPQNKVY 112

QY 61 KWEILQHLIDYILDQIALDSHPTIVSLHHQRPQGNQSRRTPLTNTDILSLQASE 120
DB 113 KWEILQHVYILDYILDQIALDSHPTIVSLHHQRPQGNQASR-LTTLNTDILSLQASE 171

QY 121 FPELSMNSDKALCG 135
DB 172 FPELSMNSDKVLCG 186

RESULT 5

I51278

negative regulatory element Idx - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 19-May-2000

C:Accession: I51278

R:Wilson, R.; Mohun, T.

Mech. Dev. 49, 211-222, 1995

A:Title: Xidx, a dominant negative regulator of bHLH function in early Xenopus embryos

A:Reference number: I51278; MUID:95252161; PMID:7734394

A:Accession: I51278

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-118 <WIL>

A:Cross-references: GB:S76880; NID:9914114; PIDN:AAB34225.1; PID:9914115

C:Genetics:

A:Gene: Xidx

A:Introns: 101/3

C:Superfamily: transcription repressor Id-2

Query Match 38.8%; Score 266.5; DB 2; Length 118;

Best Local Similarity 46.9%; Pred. No. 1.2e-18;

Matches 67; Conservative 17; Mismatches 24; Indels 35; Gaps 6;

```

Db      1  MKAISPVRSMSCYQAVCCLSQSLSIARGSSHKPGMDPEWGLLYDMNGCYSKLKEIVP 60
      52  SIPONKKVSKMEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      61  GIPQSKLSQVEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      95  GIPQSKLSQVEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      112  SILSLQASEFPPELMSNDKALC 134
      96  SILSLQKSDP-SELATQDTSVC 117

RESULT 6
I51316
negative regulatory element Id1 isoform a - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 29-Sep-1999
C:Accession: I51316
R:Zhang, H.; Reynaud, S.; Kloc, M.; Etkin, L.D.; Spohr, G.
Mech. Dev. 50, 119-130, 1995
A:Title: Id gene activity during Xenopus embryogenesis.
A:Reference number: I51316; MUID:953443988; PMID:7619724
A:Accession: I51316
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-118 <ZHA>
A:Cross-references: GB:S79007; NID:g1042004; PIDN:AAB34946.1; PID:g1042005
C:Genetics:
A:Gene: Xidfa
C:Superfamily: transcription repressor Id-2

Query Match      38.7%; Score 265.5; DB 2; Length 118;
Best Local Similarity 46.9%; Pred. No. 1.5e-18;
Matches 67; Conservative 18; Mismatches 23; Indels 35; Gaps 6;

Qy      1  MKAFSPVRSVK-----NSLSHSLGISR---KTP-VDDPMSLLYNNMDCYSKLKEIVP 51
      1  MKAISPVRSMSCYQAVCCLSQSLSIARGSSHKPGMDPEWGLLYDMNGCYSKLKEIVP 60
      52  SIPONKKVSKMEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      61  GIPQSKLSQVEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      95  GIPQSKLSQVEILQHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 111
      112  SILSLQASEFPPELMSNDKALC 134
      96  SILSLQKSDP-SELATQDTSVC 117

```

```

RESULT 7
G01855
helix-loop-helix protein Id4 - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 19-May-2000
C:Accession: G01855
R:Kiesling, T.B.
submitted to the EMBL Data Library, June 1995
A:Reference number: G08632
A:Accession: G01855
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-161 <KIE>
A:Cross-references: EMBL:U28368; NID:G881545; PIDN:AAA73923.1; PID:G881546
C:Superfamily: transcription repressor Id-2

Query Match      36.3%; Score 249; DB 2; Length 161;
Best Local Similarity 41.8%; Pred. No. 8.7e-17;
Matches 61; Conservative 17; Mismatches 32; Indels 36; Gaps 4;

Qy      1  MKAFSPVRSVKNSLS-----DHSIGIS-----RSKTPVDDP 32
      1  MKAISPVRSGRKAPSCGGGELALRCLAEHGHSLGGSAASAAAAARCAAEAADEP 60
      33  -MSLLYNNMDCYSKLKEIVSIPONKKVSKMEILQHLIDYILDQIALDSHTPTIVSL--- 88

```

```

Db      61  ALCLQCDNNDVSRRLRVPTIPPNKKVSKVEILOHVIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDI 120
      89  ----HHQRPQGNQSRRTPLTLTINTDI 110
      121  PAPPHPGAGTCPAAPPRPTLTALNTD 146

RESULT 8
A49727
DNA-binding protein inhibitor Id-1H - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Jun-2000
C:Accession: A49727
R:Hara, E.; Yamaguchi, T.; Nojima, H.; Ide, T.; Campisi, J.; Okayama, H.; Oda, K.
J. Biol. Chem. 269, 2139-2145, 1994
A:Title: Id-related genes encoding helix-loop-helix proteins are required for G-1 progr
A:Reference number: A49727; MUID:94124570; PMID:8294468
A:Accession: A49727
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-154 <HAR>
A:Cross-references: GB:D13889; NID:G464181; PIDN:BAA02988.1; PID:G471124
C:Genetics:
A:Gene: Id-1H
C:Superfamily: transcription repressor Id-2
C:Keywords: alternative splicing; cell cycle control; transcription factor

Query Match      35.3%; Score 242.5; DB 2; Length 154;
Best Local Similarity 45.0%; Pred. No. 3.5e-16;
Matches 58; Conservative 24; Mismatches 28; Indels 19; Gaps 5;

Qy      15  LSDHSLGISR-----SKTP--VDDPM--SLLYNNMDCYSKLKEIVSIPONKKVSKMEIL 65
      35  LSEQSVAISRRCGAGARLPALLDEQVNVLLYDMNGCYSKLKEIVPTLPQNRKVKVEIL 94
      66  QHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDISLSQASEFPSEL 125
      95  QHVIDYIRDQLLEINSEVGT-----PGRGPLVRAPLSTINGEISALTAEACVPA-- 147
      126  MSNDSKALC 134
      148  ---DDRILC 153

RESULT 9
S47524
gene Id1 protein - human
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 29-Sep-1999
C:Accession: S47524; S42888
R:Deed, R.W.; Jaslok, M.; Norton, J.D.
Biochim. Biophys. Acta 1219, 160-162, 1994
A:Title: Nucleotide sequence of the cDNA encoding human helix-loop-helix Id-1 protein;
A:Reference number: S47524; MUID:94368847; PMID:8086456
A:Accession: S47524
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-154 <DEE>
A:Cross-references: EMBL:X77956; NID:G457784; PIDN:CAA54920.1; PID:G457785
C:Superfamily: transcription repressor Id-2

Query Match      35.3%; Score 242.5; DB 2; Length 154;
Best Local Similarity 45.0%; Pred. No. 3.5e-16;
Matches 58; Conservative 24; Mismatches 28; Indels 19; Gaps 5;

Qy      15  LSDHSLGISR-----SKTP--VDDPM--SLLYNNMDCYSKLKEIVSIPONKKVSKMEIL 65
      35  LSEQSVAISRRCGAGARLPALLDEQVNVLLYDMNGCYSKLKEIVPTLPQNRKVKVEIL 94
      66  QHLIDYILDQIALDSHTPTIVSLHHQRPQGNQSRRTPLTLTINTDISLSQASEFPSEL 125
      95  QHVIDYIRDQLLEINSEVGT-----PGRGPLVRAPLSTINGEISALTAEACVPA-- 147

```

```
QY      126 MSNDSKALC 134
      | : |
Db      148 ---DDRILC 153

RESULT 10
JC5395
helix-loop-helix protein Id1 - human
C:Species: Homo sapiens (man)
C>Date: 04-Jun-1997 #sequence_revision 23-Aug-1997 #text_change 29-Sep-1999
C:Accession: JC5395; 165423
B:Nehlin, J.O.; Hara, E.; Kuo, W.L.; Collins, C.; Campisi, J.
Biochem. Biophys. Res. Commun. 231, 628-634, 1997
A:Title: Genomic organization, sequence, and chromosomal localization of the human helix
A:Reference number: JC5395; MUID:97224399; PMID:9070860
A:Accession: JC5395
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-155 <NEH>
A:Cross-references: GB:U57645; NID:g1816511; PIDN:AAAC13882.1; PID:g1816512
A:Experimental source: placenta cell
R:Zhu, W.; Dahmen, J.; Bulfone, A.; Rigolet, M.; Hernandez, M.C.; Kuo, W.L.; Puellies, L.
Brain Res. Mol. Brain Res. 30, 312-326, 1995
A:Title: Id gene expression during development and molecular cloning of the human Id-1 g
A:Reference number: I52661; MUID:95364622; PMID:7637581
A:Accession: I54423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-155 <RES>
A:Cross-references: GB:S78986; NID:g1042077; PIDN:AAE35037.1; PID:g1042078
C:Comment: This protein plays a role in the growth, differentiation and senescence of man
C:Genetics:
A:Gene: GDB:ID1
A:Cross-references: GDB:434745; OMIM:600349
A:Map position: 20q11-20q11
A:Introns: 142/3
C:Superfamily: transcription repressor Id-2
F:65-107/Region: helix-loop-helix #status predicted

Query Match      35.3%; Score 242; DB 2; Length 155;
Best Local Similarity 44.6%; Pred. No. 3.9e-16;
Matches 58; Conservative 24; Mismatches 28; Indels 20; Gaps 5;

QY      15 LSDHSLGTSR-----SKTP--VDDPM--SLLYNNDCYSLKELVPSIPQNKYVKWMEI 64
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      35 LSEQSVATSRGAGARLPAFLDQCVNLLYDMNGCYSLKELVPTLPQNRKYSKVEI 94

QY      65 LQHLIDYILDQLALDSHPTIVSLHHQRPQGNQSRRTTTLTNTDISILSLQASEFFPSE 124
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      95 LQHVLDYIRDLQELNSESEVGT-----PGGRGLPVRAPLSTINGEISALTAETACVPA- 148

QY      125 LMSNDSKALC 134
      | : |
Db      149 ---DDRILC 154

RESULT 11
JE0306
Id4 protein - human
C:Species: Homo sapiens (man)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: JE0306
R:Rigolet, M.; Rich, T.; Gross-Morand, M.; Molina-Gomes, D.; Viegas-Pequignot, E.; Junie
DNA Res. 5, 309-313, 1998
A:Title: cDNA cloning, tissue distribution and chromosomal localization of the human ID4
A:Reference number: JE0306; MUID:99087490; PMID:9872455
A:Accession: JE0306
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-161 <RIG>
A:Cross-references: GB:Y07958
C:Superfamily: transcription repressor Id-2
```

```
Query Match      35.3%; Score 242; DB 2; Length 161;
Best Local Similarity 41.1%; Pred. No. 4.1e-16;
Matches 60; Conservative 17; Mismatches 33; Indels 36; Gaps 4;

QY      1 MKAFSPVRSVRKNSLS-----DHSLSIGIS-----RSKTPVDDP 32
      ||| ||| | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      1 MKAVSPVRPSGRKAPSGCGGGLALRCLAEHGHSLGSGSAAAAAARCKAAEAADDP 60

QY      33 -MSLLYNNDCYSLKELVPSIPQNKYVKWMEILOHLDYILDQLALDSHPTIVSL--- 88
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      61 ALCLQCDMNDCYSLRLRLVPTIPPNKKYSKVEILQHVHDYILDQLALETHPALLRQPPP 120

QY      89 ---HHQRPQGNQSRRTTTLTNTD 110
      ||| ||| | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      121 PAPPHHPAGTCAAPPRTPLTALNTD 146

RESULT 12
S43260
helix-loop-helix protein Id4, dominant negative - mouse
C:Species: Mus musculus (house mouse)
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999
C:Accession: S43260; S37199
R:Riechmann, V.; van Cruchten, I.; Sablitzky, F.
Nucleic Acids Res. 22, 749-755, 1994
A:Title: The expression pattern of Id4, a novel dominant negative helix-loop-helix prot
A:Reference number: S43260; MUID:94188125; PMID:8139914
A:Accession: S43260
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-161 <RIE>
A:Cross-references: EMBL:X75018; NID:g402637; PIDN:CAA52926.1; PID:g402638
C:Superfamily: transcription repressor Id-2
F:64-106/Region: helix-loop-helix #status predicted

Query Match      35.1%; Score 241; DB 2; Length 161;
Best Local Similarity 42.2%; Pred. No. 5.2e-16;
Matches 62; Conservative 17; Mismatches 30; Indels 38; Gaps 5;

QY      1 MKAFSPVRSVRKNSLS-----DHSLSIGIS-----RSKTPVDDP 32
      ||| ||| | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      1 MKAVSPVRPSGRKAPSGCGGGLALRCLAEHGHSLGSGSAAAAAARCKAAEAADDP 60

QY      33 -MSLLYNNDCYSLKELVPSIPQNKYVKWMEILOHLDYILDQLALDSHPTIV--- 86
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      61 ALCLQCDMNDCYSLRLRLVPTIPPNKKYSKVEILQHVHDYILDQLALETHPALLRQPPP 120

QY      87 ---SLHHQRPQGNQSRRTTTLTNTD 110
      ||| ||| | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      121 PAPPLHPAGACPVAPPR-TPLTALNTD 146

RESULT 13
JC2111
helix-loop-helix protein, Id1 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 21-Jul-2000
C:Accession: JC2111; A42968; JC2529
R:Nagata, Y.; Todokoro, K.
Biochem. Biophys. Res. Commun. 199, 1355-1362, 1994
A:Title: Activation of helix-loop-helix proteins Id1, Id2 and Id3 during neural differ
A:Reference number: JC2111; MUID:94197727; PMID:7908517
A:Accession: JC2111
A:Molecule type: mRNA
A:Residues: 1-148 <NAG>
A:Cross-references: GB:D10862; NID:g434788; PIDN:BAA01633.1; PID:g434789
R:Springhorn, J.P.; Ellingsen, O.; Berger, H.J.; Kelly, R.A.; Smith, T.W.
J. Biol. Chem. 267, 14360-14365, 1992
A:Title: Transcriptional regulation in cardiac muscle. Coordinate expression of Id with
in vitro.
A:Reference number: A42968; MUID:92332551; PMID:1378442
A:Accession: A42968
```


A:Molecule type: mRNA
A:Residues: 1-112,'G',114-148 <SPR>
A:Cross-references: GS:M86708; NID:G203961; PIDN:AAA41090.1; PID:G203962
A:Experimental source: ventricular myocytes
A>Note: sequence extracted from NCBI backbone (NCBIN:108559, NCBIP:108559)
R:Nagata, Y.; Shoji, W.; Obinata, M.; Todokoro, K.
Biochem. Biophys. Res. Commun. 207, 916-926, 1995
A>Title: Phosphorylation of helix-loop-helix proteins Id1, Id2 and Id3.
A:Reference number: J22529; MUID:95169152; PMID:7864897
A:Accession: J22529
A:Molecule type: DNA
A:Residues: 1-148 <NA2>
A:Comment: This protein functions as an inhibitor fo cell differentiation and is a nucle
C:Comment: This protein plays a critical role in differentiation and development of a wi
C:Comment: This helix-loop-helix protein forms a heterodimer with ubiquitous and/or tis
C:Superfamily: transcription repressor Id-2
C:Keywords: differentiation; DNA binding; phosphoprotein; transcription factor
F:59-100/Region: helix-turn-helix
F:5,83/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predi
F:17/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:23/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted
F:83/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:83/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

Query Match 34.9%; Score 239.5; DB 2; Length 148;
Best Local Similarity 44.9%; Pred. No. 6.5e-16;
Matches 57; Conservative 24; Mismatches 29; Indels 17; Gaps 5;
QY 15 LSDHSLGISR---SKTP--VDDPM--SLLYNNDCYSLKELVPSIPQNKVKVKEILQH 67
DB 31 LSEQSVAISRCACTLPALLDEQQVNVLLYDMNGCYSLKELVPTLPQNRKVKVEILQH 90
QY 68 LIDYIDLQIALDSHTPTIVSLHHQRPQGNQSRRTPLTNTDILSLQASEPPELS 127
DB 91 VIDYIRDQLQELNSESEVAT-----AGARGLPVRAPLSTLNGEISALAAEACVPA---- 141

QY 128 NDSKALC 134
DB 142 -DDRILC 147

RESULT 14
A34690
DNA-binding protein inhibitor Id - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 29-Sep-1999
C:Accession: A34690
R:Benezra, R.; Davis, R.L.; Lockshon, D.; Turner, D.L.; Weintraub, H.
Cell 61, 49-59, 1990
A>Title: The protein Id: a negative regulator of helix-loop-helix DNA binding proteins.
A:Reference number: A34690; MUID:90199896; PMID:2156629
A:Accession: A34690
A:Molecule type: mRNA
A:Residues: 1-176 <BEN>
A:Cross-references: GS:M31885; NID:g194090; PIDN:AAA37879.1; PID:g194091
C:Superfamily: transcription repressor Id-2

Query Match 34.3%; Score 235.5; DB 2; Length 176;
Best Local Similarity 44.9%; Pred. No. 2e-15;
Matches 57; Conservative 24; Mismatches 29; Indels 17; Gaps 5;
QY 15 LSDHSLGISR---SKTP--VDDPM--SLLYNNDCYSLKELVPSIPQNKVKVKEILQH 67
DB 59 LSEQSVAISRCACTLPALLDEQQVNVLLYDMNGCYSLKELVPTLPQNRKVKVEILQH 118
QY 68 LIDYIDLQIALDSHTPTIVSLHHQRPQGNQSRRTPLTNTDILSLQASEPPELS 127
DB 119 VIDYIRDQLQELNSESEVGT-----TGGRGLPVRAPLSTLNGEISALAAEACVPA---- 169

QY 128 NDSKALC 134
DB 170 -DDRILC 175

RESULT 15
B49727
DNA-binding protein inhibitor Id-1H' - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Jun-2000
C:Accession: B49727
R:Hara, E.; Yamaguchi, T.; Nojima, H.; Ide, T.; Campisi, J.; Okayama, H.; Oda, K.
J. Biol. Chem. 269, 2139-2145, 1994
A>Title: Id-related genes encoding helix-loop-helix proteins are required for G-1 progr
A:Reference number: A49727; MUID:94124570; PMID:8294468
A:Accession: B49727
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-148 <HAR>
A:Cross-references: GB:D13890; NID:G464182; PIDN:BAA02989.1; PID:g471125
C:Genetics:
A:Gene: Id-1H
C:Superfamily: transcription repressor Id-2
C:Keywords: alternative splicing; cell cycle control; transcription factor

Query Match 33.1%; Score 227; DB 2; Length 148;
Best Local Similarity 47.3%; Pred. No. 1e-14;
Matches 53; Conservative 21; Mismatches 24; Indels 14; Gaps 4;
QY 15 LSDHSLGISR-----SKTP--VDDPM--SLLYNNDCYSLKELVPSIPQNKVKVKEIL 65
DB 35 LSEQSVAISRCACTLPALLDEQQVNVLLYDMNGCYSLKELVPTLPQNRKVKVEIL 94
QY 66 QHLIDYIDLQIALDSHTPTIVSLHHQRPQGNQSRRTPLTNTDILSLQ 117
DB 95 QHVIDYIRDQLQELNSESEVGT-----PGRGLPVRAPLSTLNGEISALTAE 141

Search completed: September 27, 2004, 11:28:46
Job time : 18 secs

Query Match	35.3%	Score 242;	DB 15;	Length 155;
Best Local Similarity	44.6%;	Pred. No. 1.6e-18;		
Matches	58;	Conservative 24;	Mismatches 28;	Indels 20; Gaps 5;
QY	15	LSDSHSLGISR-----SKTP--VDDPM--SLIYNMDCYSLKELVPSIQNKVKSVKEI	64	
Db	35	LSQSVAISRCAGGAGARLFDLDEQVNVLLYDMNGCYSLKELVPTLPQNRKVKSVKEI	94	
QY	65	LGHLIDVYLDLQALDQSHPTIVSLHHQRPQCNQSRITPLTITNTDLSILQASERPSE	124	
Db	95	LQHVIDIYRDQLQELNSESVEGTF-----PGGRGLPVRAPLSTLNGEISALTAAACVPA-	148	

```

RESULT 8
US-10-102-806-722
; Sequence 722, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 722
; LENGTH: 111

```

Query Match 34.5%; Score 236.5; DB 14; Length 111;

```

US-09-897-787-20
Query Match          30.0%; Score 212; DB 12; Length 129;
Best Local Similarity 44.2%; Pred.No. 2.5e-15;
Matches 50; Conservative 21; Mismatches 28; Indels 14; Gaps 4;

QY      1 MKAFSPVRVRKN--SLSDHSLSGIS--RSKTP-VDDPMSLLYNMDCYSKLKELVPSIPQ 55
Db       1 MKALSPVRGCEAVCCLSERSLAIARGRGKSAAEPLSLDDNMHCYSRLRELVPGVPR 60
           ::::|||||:|||||||:-----DSHPTTVSLHHOREPGONQRS 99
           ::::::|||:|||:|||||:
Db       61 GTQLSQVEILQRVIDYLIDLVQVLAEPAFGPPDPGHLPQVREGARPGSSERA 113

RESULT 11
US-10-324-618-16
Sequence 16, Application US/10324618
Publication No. US20030170691A1
GENERAL INFORMATION:
APPLICANT: Gimeno, Ruth
APPLICANT: Wu, Zhidan
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Hubbard Brian K.
TITLE OF INVENTION: HUMAN DIACYLGlycerol ACYLTRANSFERASE 2
FILE REFERENCE: MPI01-263P2RM
CURRENT APPLICATION NUMBER: US/10/324,618
CURRENT FILING DATE: 2002-12-19
PRIOR APPLICATION NUMBER: 60/341,947
PRIOR FILING DATE: 2002-12-19
PRIOR APPLICATION NUMBER: 60/411,859
PRIOR FILING DATE: 2002-09-19
NUMBER OF SEQ ID NOS: 65
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 374
TYPE: PRT
ORGANISM: mus musculus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(470)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-324-618-16

Query Match          30.0%; Score 206; DB 14; Length 374;
Best Local Similarity 86.4%; Pred.No. 5e-14;
Matches 38; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      17 DHSIGISRSKTVPDPMSSLNMDNCYSKLKELVPSIPQNKKVS 60
Db       322 DHSLGISQSQTPEVDPMSSLNMDNCYSKLKELMFESIPQNKAA 365
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-925-298-457
Sequence 457, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103
CURRENT APPLICATION NUMBER: US/09/925,298
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 457
LENGTH: 152
TYPE: PRT
```


Db 44 SKLXDLVFPMPKXNRKLTKEIIQHVIDYICDLQTELETHPEMGNFDAAAAALTAVNGLHED 103
Qy 92 RPGNQSRSTTTLTNTDTSILSLQASEFPSELMS 127
Db 104 EDSMEDADAEAEVDPDLAQLNA-EQPAKVSS 138

Search completed: September 27, 2004, 11:39:08
Job time : 132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 11:26:27 ; Search time 19 seconds
(without alignments)
366.816 Million cell updates/sec

Title: US-10-025-170-2
Perfect score: 686
Sequence: 1 MKAFSPVRKNSLDHSL.....LQAEFPSELNDSKALCG 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	661.5	96.4	134	3	US-09-045-764A-8
2	654.5	95.4	134	3	US-09-045-764A-9
3	646.5	94.2	134	3	US-09-045-764A-2
4	640	93.3	133	3	US-09-045-764A-11
5	522	76.1	135	3	US-09-045-764A-10
6	266.5	38.8	128	3	US-09-045-764A-5
7	252.5	36.8	128	3	US-09-045-764A-6
8	242.5	35.3	154	1	US-08-151-391A-2
9	241	35.1	161	3	US-09-045-764A-4
10	235.5	34.3	148	3	US-09-045-764A-1
11	227	33.1	148	1	US-08-151-391A-4
12	217	31.6	119	1	US-08-256-077-4
13	217	31.6	119	1	US-08-466-127-4
14	217	31.6	148	1	US-08-256-077-2
15	217	31.6	148	1	US-08-466-127-2
16	208	30.3	119	3	US-09-045-764A-3
17	186	27.1	46	1	US-08-466-127-8
18	152	22.2	36	3	US-09-045-764A-12
19	148	21.6	46	1	US-08-466-127-7
20	139.5	20.3	199	3	US-09-045-764A-7
21	128	18.7	46	1	US-08-466-127-9
22	94.5	13.8	214	4	US-08-722-570-2
23	94.5	13.5	214	4	US-08-932-411A-2
24	92.5	13.5	244	4	US-08-910-973-22
25	92.5	13.5	244	4	US-09-499-227-22
26	92.5	13.5	244	4	US-08-932-411A-14
27	91.5	13.3	237	1	US-08-910-973-13

28 91.5 13.3 237 4 US-09-499-227-13 Sequence 13, Appli
29 91.5 13.3 244 4 US-08-722-570-1 Sequence 1, Appli
30 91.5 13.3 244 4 US-08-932-411A-1 Sequence 1, Appli
31 88 12.8 238 3 US-09-234-332-10 Sequence 10, Appli
32 88 12.8 238 4 US-09-702-705-1813 Sequence 1813, Ap
33 88 12.8 238 4 US-09-736-457-1813 Sequence 1813, Ap
34 88 12.8 238 4 US-09-671-325-1813 Sequence 4, Appli
35 83 12.1 57 4 US-08-722-570-4 Sequence 4, Appli
36 83 12.1 57 4 US-08-932-411A-4 Sequence 20, Appli
37 75 10.9 214 4 US-09-914-259-12 Sequence 12, Appli
38 75 10.9 1087 4 US-08-620-694A-2 Sequence 2, Appli
39 74.5 10.9 864 2 US-09-022-255-2 Sequence 2, Appli
40 74.5 10.9 864 3 US-09-022-255-2 Sequence 2, Appli
41 74.5 10.9 864 3 US-08-978-773-2 Sequence 2, Appli
42 74.5 10.9 864 3 US-09-022-253-2 Sequence 2, Appli
43 74.5 10.9 864 3 US-09-022-260-2 Sequence 2, Appli
44 74.5 10.9 864 3 US-09-022-259-2 Sequence 2, Appli
45 74.5 10.9 864 3 US-09-022-259-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-045-764A-8
; Sequence 9, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-045-764A-8

Query Match 96.4%; Score 661.5; DB 3; Length 134;
Best Local Similarity 97.8%; Pred. No. 4.3e-72;
Matches 132; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKAFSPVRKNSLDHSLGIGISRSKTPVDDPMSLLYNNDCVSKLKEIVPSIPQNKVYS 60
|||||

Db 1 MKAFSPVRKNSLDHSLGIGISRSKTPVDDPMSLLYNNDCVSKLKEIVPSIPQNKVYS 60
|||||

QY 61 KMEILQHLIDYILDQLDIALDSHTPTVSLHHQPGNQSRRTTFLTINTDITSLQASE 120

RESULT 3
US-09-043-764A-2
; Sequence 2, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica

GENERAL INFORMATION:
APPLICANT: Israel, Mark A.
APPLICANT: Florio, Monica
TITLE OF INVENTION: Apoptotic Peptides
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-045-764A-11

Query Match      93.3%; Score 640; DB 3; Length 133;
Best Local Similarity 96.3%; Pred. No. 1.7e-69;
Matches 130; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY      1 MKAFSPVRSVRKNSLDHSLGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIPQNKVS 60
Db      1 MKAFSPVRSVRKNSLDHSLGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIPQNKVS 60

QY      61 KWEILQHLIDYILDQLALDHSPTIVSLHHQRPQGNQSRRTPLTLNTDISLSLQASE 120
Db      61 KWEILQHLIDYILDQLALDHSPTIVSLHHQRPQGNQASR-TPL-TLNTDISLSLQASE 118

QY      121 FPELSMNSDKALCG 135
Db      119 FPELSMNSDKALCG 133

RESULT 5
US-09-045-764A-10
; Sequence 10, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-045-764A-5

Query Match      38.8%; Score 266.5; DB 3; Length 118;
Best Local Similarity 46.9%; Pred. No. 1.7e-24;
Matches 67; Conservative 17; Mismatches 24; Indels 35; Gaps 6;

QY      1 MKAFSPVRSVRK-----NSLDHSLGSR-----KTP-VDDPMSLLYNNMDCYSKIKELVP 51
Db      1 MKAFSPVRSVRKNSLDHSLGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIPQNK 60

QY      52 SIPQNKVKWEILQHLIDYILDQLALDHSPTIVSLHHQRPQGNQSRRTPLTLNTDI 111
Db      61 GIPQSKLSQVEILQHLIDYIFDLQIVL-----GEDQQQS----- 95

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-045-764A-10

Query Match      76.1%; Score 522; DB 3; Length 135;
Best Local Similarity 79.3%; Pred. No. 3e-55;
Matches 107; Conservative 11; Mismatches 13; Indels 4; Gaps 2;

QY      1 MKAFSPVRSVRKNS--LSDHSLGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIPQNK 58
Db      1 MKAFSPVRSVRKNSNLSEHSLGSRKTPVDDPLSLYNNMDCYSKIKELVPSIPQNK 60

QY      59 VSKWEILQHLIDYILDQLALDHSPTIVSLHHQRPQGNQSRRTPLTLNTDISLSLQ 118
Db      61 VSKWEILQHLIDYILDQLALDHSNVAITSHHPRPG--QATPRTPLTLNTDISLSLQ 118

QY      119 SEFPSELSMNSDKAL 133
Db      119 SEFPSELSMNSDKAL 133

RESULT 6
US-09-045-764A-5
; Sequence 5, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-045-764A-5

Query Match      38.8%; Score 266.5; DB 3; Length 118;
Best Local Similarity 46.9%; Pred. No. 1.7e-24;
Matches 67; Conservative 17; Mismatches 24; Indels 35; Gaps 6;

QY      1 MKAFSPVRSVRK-----NSLDHSLGSR-----KTP-VDDPMSLLYNNMDCYSKIKELVP 51
Db      1 MKAFSPVRSVRKNSLDHSLGSRKTPVDDPMSLLYNNMDCYSKIKELVPSIPQNK 60

QY      52 SIPQNKVKWEILQHLIDYILDQLALDHSPTIVSLHHQRPQGNQSRRTPLTLNTDI 111
Db      61 GIPQSKLSQVEILQHLIDYIFDLQIVL-----GEDQQQS----- 95
```

```
QY 112 SILSLQASEFPSELSMNSDKALC 134
    ||||| | | | | : |
Db 96 SILSLQKSDF-SELATQGTISVC 117

RESULT 7
US-09-045-764A-6
; Sequence 6, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-045-764A-6

Query Match 36.8%; Score 252.5; DB 3; Length 128;
Best Local Similarity 41.8%; Pred. No. 9.6e-23;
Matches 61; Conservative 24; Mismatches 30; Indels 31; Gaps 6;

QY 1 MKAFSPVRSVRKNS-----LSDHSLGTSRSTTP-VDDPMSL-LYNNMDCYSKSLKEL 49
    ||| : : :
Db 1 MKVWGPTCALRSKVGGEVDWRCLSDQSLAISKCKIPLDDEQMTWFLQDMNCSYKSLKEL 60
    ||| : : :

QY 50 VPSIQNKVKSWKWEILQHLIDYLDQALDSHPITVSLHHQRPQGNQSR-TPPLTTLN 108
    ||| : : :
Db 61 VPTHPTNKASKWEILQHVLDYLDQALDSHPITVSLHHQRPQGNQSR-TPPLTTLN 108
    ||| : : :

QY 109 TDISLSLQASEFPSELSMNSDKALC 134
    ||| : : :
Db 109 AELASISVNG-----CSDDRIMC 127

RESULT 8
US-08-151-391A-2
; Sequence 2, Application US/08151391A
; Patent No. 5527897
; GENERAL INFORMATION:
; APPLICANT: Oda, Kinichiro
; APPLICANT: Nakada, Susumu
; APPLICANT: Hara, Eiichi
; APPLICANT: Yamaguchi, Tomoko

; APPLICANT: Nakamura, Takeshi
; APPLICANT: Oka, Yumiko
; APPLICANT: Kishimoto, Toshihiko
; TITLE OF INVENTION: Human ID Genes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,391A
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mills, Demetra J.
; REGISTRATION NUMBER: 34,506
; REFERENCE/DOCKET NUMBER: 715-087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)684-1111
; TELEFAX: (703)684-1124
; TELEX: 82-4425
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 154 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-151-391A-2

Query Match 35.3%; Score 242.5; DB 1; Length 154;
Best Local Similarity 45.0%; Pred. No. 2e-21;
Matches 58; Conservative 24; Mismatches 28; Indels 19; Gaps 5;

QY 15 LSDHSLGTSR-----SKTP--VDDPM--SLYNNMDCYSKSLKELVPSIQNKVKSWKWEIL 65
    ||| : : :
Db 35 LSEQSVALSRCRGAGARLPALLDEQVNVLLYDNGCYSLKELVPTLPQNKVKSWKWEIL 94
    ||| : : :

QY 66 QHLIDYLDQALDSHPITVSLHHQRPQGNQSR-TPPLTTNTDISLSLQASEFPSEL 125
    ||| : : :
Db 95 QHVLDYLDQALDSHPITVSLHHQRPQGNQSR-TPPLTTNTDISLSLQASEFPSEL 147
    ||| : : :

QY 126 MSNDSKALC 134
    ||| : : :
Db 148 ---DDRILC 153

RESULT 9
US-09-045-764A-4
; Sequence 4, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
```

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/045,764A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OSMAN, RICHARD A
;; REGISTRATION NUMBER: 36,627
;; REFERENCE/DOCKET NUMBER: UCSF98-045
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650) 343-4341
;; TELEFAX: (650) 343-4342
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 161 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-045-764A-4

Query Match 35.1%; Score 241; DB 3; Length 161;
Best Local Similarity 42.2%; Pred. No. 3.3e-21;
Matches 62; Conservative 17; Mismatches 30; Indels 38; Gaps 5;
QY 1 MKAFSPVRSVRKNSLS-----DHSIGIS-----RSKTPVDDP 32
Db 1 MKAVSPVRSRKAPSGGGGELALRCIAEHGSLGSSAAAAAARCAAEAADEP 60
QY 33 -MSLLVNMDCYSLKELVSPQNKKVKMEILQHLIDYLDLQIALDSHPTIV----- 86
Db 61 ALCLQCMDCYSLRLRLVPTIPNKKVKVEILQHLIDYLDLQIALETHPALLRPPPP 120
QY 87 ---SLHHQPGQNRSTPTLTNTD 110
Db 121 PAPLHPAGACFVAPPR-TPLTALNTD 146

RESULT 10
US-09-045-764A-1
; Sequence 1, Application US/09045764A
; Patent No. 6127178
; GENERAL INFORMATION:
; APPLICANT: Israel, Mark A.
; APPLICANT: Florio, Monica
; TITLE OF INVENTION: Apoptotic Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,764A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCSF98-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342

;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 148 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-045-764A-1

Query Match 34.3%; Score 235.5; DB 3; Length 148;
Best Local Similarity 44.9%; Pred. No. 1.3e-20;
Matches 57; Conservative 24; Mismatches 29; Indels 17; Gaps 5;
QY 15 LSDHSLGISR---SKTP-VDPDM--SLLYNMDCYSLKELVSPQNKKVKMEILQHL 67
Db 31 LSEQVAISRCASTRLPALDDEQVNVLLYDMGYSRLKELVPTLPQNRKVKVEILQHL 90
QY 68 LIDYLDLQIALDSHPTIVSLHHQPGQNRSTPTLTNTDIDISILQASEPPELS 127
Db 91 VIDIRDLQELNSESEVGT-----TGGRLPVVRAPLSTLNGEISALAAEAACVPA---- 141
QY 128 NDSKALC 134
Db 142 -DDRILC 147

RESULT 11
US-08-151-391A-4
; Sequence 4, Application US/08151391A
; Patent No. 5527897
; GENERAL INFORMATION:
; APPLICANT: Oda, Kinichiro
; APPLICANT: Nakada, Susumu
; APPLICANT: Hara, Eiichi
; APPLICANT: Yamaguchi, Tomoko
; APPLICANT: Nakamura, Takeshi
; APPLICANT: Oka, Yumiko
; APPLICANT: Kishimoto, Toshihiko
; TITLE OF INVENTION: Human ID Genes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,391A
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mills, Demetra J.
; REGISTRATION NUMBER: 34,506
; REFERENCE/DOCKET NUMBER: 715-087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 684-1111
; TELEFAX: (703) 684-1124
; TELEX: 82-4425
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-151-391A-4

[illegible]

RESULT 13
US-08-466-127-4

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,077
FILING DATE: 23-JUNE-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.1360000
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 148 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-077-2

Query Match 31.6%; Score 217; DB 1; Length 148;
Best Local Similarity 39.3%; Pred. No. 2.3e-18;
Matches 55; Conservative 24; Mismatches 33; Indels 28; Gaps 6;
QY 1 MKAFSPVRSVRKN--SLSDHSLGSRSK---TPVDDPMSLLYNNMDCYSKLKELVPSIPQ 55
Db 30 MKALSPVRCGYEAVCCLSERSLAIRGKGPAABEPLSLDDMMHCYSRLRELVPVGR 89
QY 56 NKKVSKMELLQHLIDYILDQIALDSHPTIVSLHHQRPQGNQRSTPLTLTNTDISLS 115
Db 90 GTQLSQVEILLQRLVIDYILDQVVL-ABPA-----PGP-----PDGPHLP 127
QY 116 LOASEFPESEL-MSNDKALC 134
Db 128 IQTAELAPELVISNDRSFC 147

RESULT 15
US-08-466-127-2
Sequence 2, Application US/08466127
Patent No. 5683878
GENERAL INFORMATION:
APPLICANT: Ellmeier, Willfried
APPLICANT: Weith, Andreas
TITLE OF INVENTION: Neuroblastoma-Associated Regulator Gene
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,127
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.1360002
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 148 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-127-2
Query Match 31.6%; Score 217; DB 1; Length 148;
Best Local Similarity 39.3%; Pred. No. 2.3e-18;
Matches 55; Conservative 24; Mismatches 33; Indels 28; Gaps 6;
QY 1 MKAFSPVRSVRKN--SLSDHSLGSRSK---TPVDDPMSLLYNNMDCYSKLKELVPSIPQ 55
Db 30 MKALSPVRCGYEAVCCLSERSLAIRGKGPAABEPLSLDDMMHCYSRLRELVPVGR 89
QY 56 NKKVSKMELLQHLIDYILDQIALDSHPTIVSLHHQRPQGNQRSTPLTLTNTDISLS 115
Db 90 GTQLSQVEILLQRLVIDYILDQVVL-ABPA-----PGP-----PDGPHLP 127
QY 116 LOASEFPESEL-MSNDKALC 134
Db 128 IQTAELAPELVISNDRSFC 147
Search completed: September 27, 2004, 11:29.10
Job time : 20 secs

B/GMK